

Financial Constraints and Girls' Secondary Education

Evidence from School Fee Elimination in The Gambia

Moussa P. Blimpo

Ousman Gajigo

Todd Pugatch



WORLD BANK GROUP

Africa Region

Office of the Chief Economist

December 2016

Abstract

This study analyzes the impact of large-scale fee elimination for secondary school girls in The Gambia on the quantity, composition, and achievement of students. The gradual rollout of the program across geographic regions provides identifying variation in the policy. The program increased the number of girls taking the high school exit exam by 55 percent. The share of older test takers increased in poorer districts, expanding access for students who began school late, repeated grades, or whose studies had been interrupted. Despite these changes in the quantity and composition of

students, there are robustly positive point estimates of the program on test scores, with suggestive evidence of gains for several subgroups of both girls and boys. Absence of learning declines is notable in a setting where expanded access could strain limited resources and reduce school quality. The findings suggest that financial constraints remain serious barriers to post-primary education, and that efforts to expand access to secondary education need not come at the expense of learning in low-income countries like The Gambia.

This paper is a product of the Office of the Chief Economist, Africa Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at mblimpo@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Financial Constraints and Girls' Secondary Education: Evidence from School Fee Elimination in The Gambia*

Moussa P. Blimpo, Ousman Gajigo, and Todd Pugatch

Keywords: Gambia, gender gap, school fee elimination, secondary school
JEL codes: I21, I25, I28, O15.

Most countries in Sub-Saharan Africa experienced large expansions of access to primary education over the past two decades. For example, the number of primary school children doubled between 1998 and 2009 in countries like Burkina Faso, Madagascar, Mali, and Mozambique (World Bank 2016). Despite such success at the primary level, gross secondary enrollment remains low in the region, at 46% for boys and 39% for girls in 2013 (World Bank 2015). Large gender enrollment gaps in many Sub-Saharan African countries pose additional challenges for girls seeking to pursue their education beyond the primary grades. One potential explanation for low secondary enrollment is financial

Moussa P. Blimpo is an Economist in the Office of the Chief Economist for the Africa Region at the World Bank; his email address is mblimpo@worldbank.org. Ousman Gajigo is a consultant for the World Bank; his email address is ousman_g@yahoo.com. Todd Pugatch (corresponding author) is an Assistant Professor

constraints. Relative to primary school, the overall cost of attending secondary school is much larger due to higher tuition fees, higher opportunity costs as children are older and may earn more on the labor market, and transport costs associated with fewer secondary school choices, especially in rural areas. To date, only a handful of countries in the region, such as South Africa, Ghana, and The Gambia, offer large-scale tuition-free secondary education or some form of financial aid through various scholarship programs.

The Gambia has been a pioneer in promoting access to secondary education, offering fee-free public schooling for girls in grades 7–12 on a nearly national scale for more than a decade. In this paper, we evaluate this policy, known as the girls’ scholarship program, on student learning. Two features of the policy make it especially suited for evaluation. First, the program was rolled out to different regions on a staggered schedule between 2001 and 2004. This allows us to use the regions that received the program later as a control group, exploiting variation in program receipt over time and across regions. Second, the program exclusively targeted girls, allowing us to measure whether boys within program regions experienced spillover effects from the program. To our knowledge, this is the first paper to evaluate the impact of a large-scale tuition waiver program at the secondary school level in Africa.¹

We find that the program had important effects on both school access and student achievement. The policy increased the quantity of girls taking the high school exit exam by 26.1 students per district, or 55%. The share of girls enrolled in grade 12 attempting the exam increased by 28 percentage points, bringing more students near secondary school completion. The share of older test takers increased in poorer districts, expanding access for students who began school late, repeated grades, or whose studies had been interrupted. Despite these changes in the quantity and composition of students, we find robustly positive point estimates of the program on test scores, with suggestive evidence of gains for several student subgroups. The program improved education outcomes for boys as well, with qualitatively similar increases in access and achievement. Enrollment spillovers for boys were concentrated among households with older girls who benefitted from the program, consistent with alleviation of household financial constraints. In light of these results, we conclude that the program expanded access without harming learning outcomes.

These findings are notable because expanded access might strain limited resources or reduce the average quality of students or schools. In the words of (Banerjee et al. 2007, 1236), “Ironically, the difficulty in improving the quality of education may in part be a by-product of the success in getting more children to attend school.” Yet this tradeoff may not be as stark in secondary schools starting from a low enrollment base, as in this study. Moreover, in recent years the Gambian government has engaged several initiatives toward improving learning outcomes, including promoting decentralized school governance (Blimpo and Evans 2011) and salary premiums for teachers in rural schools (Pugatch and Schroeder 2014a, 2014b). The findings from this study suggest that complementary efforts to expand access to schooling, such as the girls’ scholarship program, need not impede school quality. Moreover, programs to lower schooling costs for girls can assist boys as well by removing household financial constraints.

As mentioned previously, the study occurs against a backdrop of sustained attention among both policy makers and researchers to primary education, with relatively less emphasis on the secondary level. Expanded access to primary education resulted from concerted policies, both internationally and nationally, aimed at removing financial constraints through school fee elimination and other measures. Enrollments in primary schools have accelerated in many countries since the 1990 Jomtien conference, in which over 150 countries adopted the Education For All initiative. This commitment was renewed during the Dakar Framework for Action in 2000, which targeted tuition elimination and other cost reductions. Over the past two decades, more than twenty African countries have waived tuition from primary education and many more have some form of targeted programs to ease access to the most disadvantaged populations. Several recent literature reviews concluded that the great majority of interventions that reduced tuition fees and other costs increased enrollment, suggesting that financial constraints are among the most important barriers to access to primary education in poorer developing countries (Petrosino et al. 2012; Krishnaratne et al. 2013; Murnane and Ganimian 2014).

These great successes on access have been achieved amid growing concerns about education quality and potential degradation of learning outcomes (Pritchett 2013). More recent research has focused on the impact of access-oriented policies not only on enrollment but also on student performance. For example, Kazianga et al. (2013) found that a comprehensive program in Burkina Faso that included school construction and student attendance incentives increased enrollment and test scores of primary school students. A similar, but experimental, study that brought community schools to Afghan villagers found equally large effects on both enrollment and test scores (Burde and Linden 2013). In Kenya, Lucas and Mbiti (2012a) found that elimination of primary school fees led to substantial enrollment gains with little negative effect on the test scores of those who would have attended in the absence of the tuition waiver. These studies suggest that at least for primary schools, the tradeoff between access and learning might be less pronounced than one might think.

Given these successes in improving access and (to a lesser extent) learning in primary education, for many countries the logical next step is to improve access and outcomes in secondary education. Fewer policies and studies have focused on secondary schools, however. Early results from an ongoing study on a scholarship program in Ghana found large enrollment effects among scholarship winners relative to the control group three years after the program started (Duflo et al. 2009). They concluded that financial barriers might be crucial at the secondary level as well. Outside of Africa, Muralidharan and Prakash (2013) evaluated a program in India that reduced girls' cost of attending secondary school through provision of bicycles, increasing enrollment by 30% and cutting the gender gap by 40%. Yet major gaps in understanding remain, particularly with regard to student achievement. A review of the post-primary schooling literature by Banerjee et al. (2013) concluded, "Despite the overarching positive results of price-based policies in increasing school enrollment and attendance, the evidence on the effects of price reductions on student performance is less conclusive" (21).

We contribute to that literature by being the first to evaluate the achievement effects of a large-scale tuition elimination policy for secondary education in Africa. The present study extends work by one of this study's authors (Gajigo 2014), who used household survey data to find large enrollment gains from the same program. We use administrative data on the universe of standardized test scores in The Gambia from 1998 to 2012; we are the first researchers to obtain and analyze this data. We also contribute to the broader literature on efforts to close the gender gap in access and learning. Several other studies have evaluated similar programs targeting girls (Kim et al. 1999a, 1999b, and Chaudhury and Parajuli 2010 for Pakistan; Filmer and Schady 2008 for Cambodia; Kremer et al. 2009 for Kenya; Baird et al. 2011 for Malawi; Begum et al. 2012 for Bangladesh; and the previously mentioned Kazianga et al. 2013 and Muralidharan and Prakash 2013 for Burkina Faso and India, respectively), with a consensus finding that reducing the cost of attendance leads to gains in enrollment. Of these, however, only Baird et al. (2011) examines learning outcomes among secondary school students as we do, using a program more local in scope than our setting.

In the next section, we describe the education system in The Gambia and the girls' scholarship program. Sections III–IV present the methodology and data we use for analysis. Section V presents results, and Section VI concludes.

I. EDUCATION IN THE GAMBIA AND THE GIRLS' SCHOLARSHIP PROGRAM

In the Gambian education system, the first 9 years are formally known as the Basic Cycle. This includes six years of primary school (grades 1–6) and three years of Upper Basic School (middle school, grades 7–9). High school, known locally as Senior Secondary School, consists of grades 10–12. The West African Senior School Certificate Examination (WASSCE, hereafter the Grade 12 exam), instituted in 1998, is administered at the end of grade 12 and is required for advancing to university. The West African Examination Council (WAEC), a regional institution that conducts examinations in the four former British colonies in West Africa (The Gambia, Ghana, Nigeria, and Sierra Leone), administers the exam.

WAEC generates exam questions each year in consultation with the Ministry of Education, based on existing curricula. Accordingly, the exam measures achievement in specific subjects, rather than innate ability. Students choose a minimum of six and a maximum of nine subjects, but the core and mandatory subjects are mathematics and English, which will be our focus. There is no fixed passing mark for the exam. Because the exam is based on curricula designed by the Ministry of Education, and these have not undergone any major change, the exam questions should be comparable over time. Students must register and pay a fee of approximately US\$30 to take the exam.

The exam takes place within a structured system. Each year, sealed questions are delivered at the test centers the day before the scheduled exam.² On the day of the exam, teachers from other schools serve as invigilators (proctors). The exams are centrally graded by WAEC. This structure is similar to the way national exams are conducted in

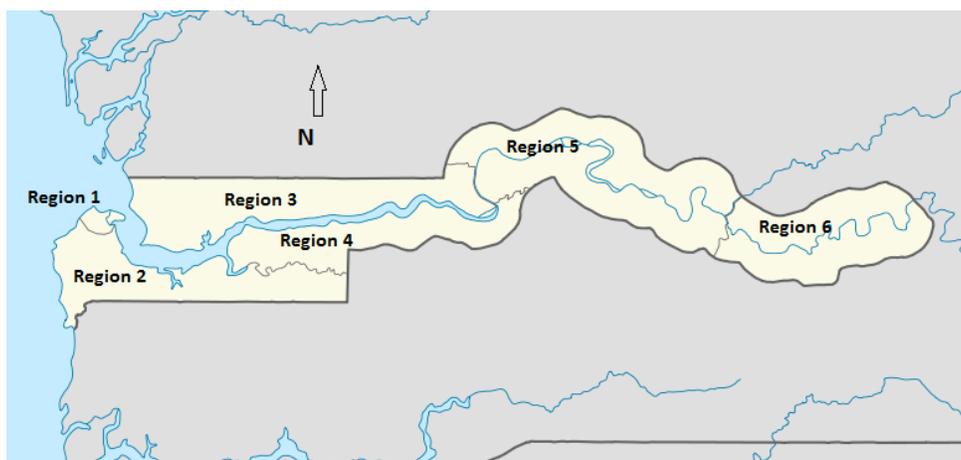
other countries (Kremer et al. 2009). Exams are high-stakes for students, because results are used in university admissions and public sector hiring. However, school resources and teacher salaries are not tied to student performance, mitigating concerns that schools might discourage weaker students from taking the exam as in other settings (Cullen and Reback 2006).

Like other African countries, The Gambia charged fees for public school attendance until last year. The Gambia levied fees beginning in grade 7, as primary education is nominally free for public schools since 2013. Students are still responsible for purchasing textbooks, uniforms, and other materials, leading students to bear costs even at the primary school level.¹

The scholarship program for female middle and high school students started as an initiative funded jointly by UNICEF, the World Bank, and the International Monetary Fund through the Highly Indebted Poor Countries program and the Gambian government. The goal of the program is to increase overall student enrollment but with a specific focus on reducing the gender gap. The program pays mandatory school fees, including exam fees, for all girls in grades 7–12 in the regions in which it is implemented.³ The only criteria for benefitting from the program are gender (female) and attending a public secondary school.⁴

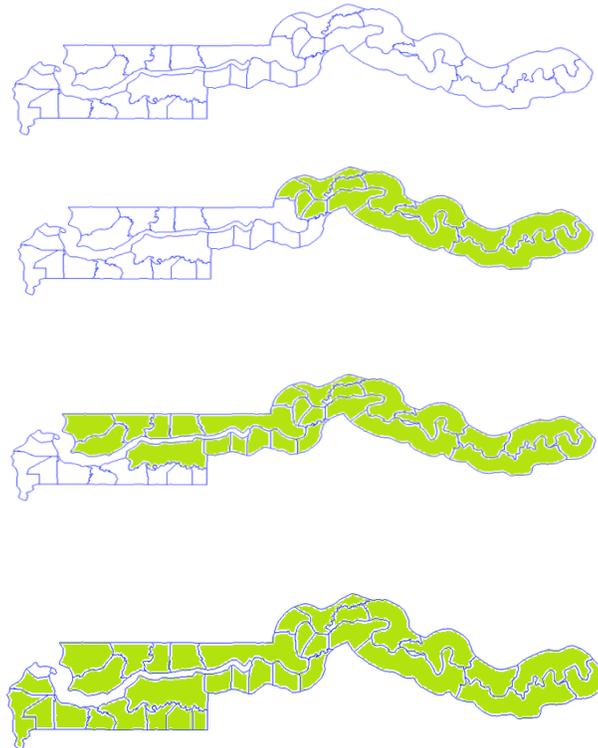
The scholarship program began in 2001 in regions 5 and 6 only, as these are the regions that are most rural and have the lowest enrollment.⁵ The program was extended to regions 3 and 4 the following year. Two academic years later in 2004, the program expanded further to include region 2. The scholarship program was not extended to region 1, the most urbanized and developed region, until 2014, two years after the end of our sample. Figure 1 provides a map of The Gambia's regions, while figure 2 shows the rollout of the program over time.

Figure 1: The Gambia and its regions



¹ The Gambia eliminated fees in Upper Basic Schools in 2014 and Senior Secondary Schools in 2015. The government now also plans to provide textbooks. Earlier draft of this work did not include this information as it was not in effect yet.

Figure 2(a)-(d): Girls' Scholarship Program implementation



To implement the program, a specially designated Ministry of Education administrator handles the disbursement of funds between the program and schools. Transfers are based on the number of girls enrolled per school. The regional offices of the Ministry verify the enrollment figures provided by individual schools before the scholarship funds are transferred. At no point do the beneficiary households handle the money, thereby removing any chance of the scholarship funds being diverted for other purposes. The average cost of the program per student was US\$48, US\$43, US\$42, and US\$43 in 2001, 2002, 2003, and 2004, respectively (of Basic and Secondary Education, 2004).⁶ The benefit is particularly large in grades 10–12, where fees are more than 7 times those for grades 7–9 (Daly et al. 2014). The program was revenue neutral for schools on a per student basis because girls were previously charged fees equivalent to the scholarship.⁷ The program was widely publicized through local media, as well as through several workshops in various regions of the country. No other policy coincided with the scholarship program in geographic scope and timing.

II. METHODOLOGY

This paper analyzes the effect of the Gambian girls' scholarship program on student access and learning, using the geographically staggered rollout of the program to compare outcomes in regions that received the program early with those that received it

late. Additionally, comparing results for girls and boys within regions tests whether targeting the program to girls led to differential effects by gender.

We use a difference-in-differences identification strategy to evaluate the program. We estimate the following regression separately for boys and girls:

$$y_{isrt} = \beta D_{rt} + X_{isrt}\gamma + \delta_s + \theta_t + \varepsilon_{isrt} \quad (1)$$

where y_{isrt} is the outcome (i.e., test score) of student i at school s in region r in year t ; D_{rt} is a dummy for whether the scholarship program was implemented in region r at time t ; X is a vector of individual characteristics, including the student's age (measured continuously, based on date of birth), age squared, and a constant; and δ and θ are school and time fixed effects, respectively. The coefficient β is the difference-in-differences estimate of the effect of the program because it compares changes in test performance of students in regions that had received the program by time t to changes in regions that had not.

The identifying assumption is that in the absence of the program, changes in outcomes in regions that received the program early would have been the same as in regions that received the program late. We examine the validity of this assumption by testing for common pretreatment trends across regions. To do so, we rescale time so that $t = 0$ corresponds to the year of treatment receipt in each region and limit the sample to pre-treatment periods only. We then regress outcomes on a time trend and its interaction with indicators for regions 5 and 6 (which received the program first, in 2001) and regions 3 and 4 (which received the program in 2002):

$$y_{isrt} = \alpha_0 \tilde{t} + \alpha_1 \tilde{t} * \mathbf{1}(\text{Region} = 3,4) + \alpha_2 \tilde{t} * \mathbf{1}(\text{Region} = 5,6) + X_{isrt}\gamma + \delta_s + \theta_t + \varepsilon_{isrt} \quad (2)$$

where \tilde{t} is the rescaled time trend and all else is as in equation (1). Region 2 is the omitted category because we drop region 1 (Banjul, the capital) from all analysis due to its dissimilarity with the rest of the country. Statistically significant coefficients on the interaction terms would indicate differential pre-treatment trends among regions, calling into question the identifying assumption of our difference-in-differences strategy.

Even if the identifying assumption holds, proper interpretation of the parameter of interest β in (1) bears reflection. The goal of the policy was to increase secondary school access for girls, which if successful would alter the quantity and composition of students taking the test. Students induced to enroll by the program are likely to be less

academically prepared than their peers for whom financial barriers are not a constraint. Additionally, an influx of students could strain school resources. Each of these channels would lead to a negative effect of the policy on learning. On the other hand, relaxing financial constraints among students who would have enrolled in the absence of the policy could improve learning by reducing their need to generate income or by alleviating stress.

The treatment effect β will therefore represent an average of these effects along the extensive and intensive margins, or what Glewwe and Muralidharan (2015) call the “policy parameter,” because it represents the effect of the policy inclusive of any adjustments made by households or schools in response. We therefore analyze how the policy altered the number of students taking the exam, the share of test takers in scholarship-eligible schools, and student characteristics, in order to understand selection into the test. We also look for heterogeneity in treatment effects by interacting the program dummy with observable characteristics, allowing us to check whether treatment effects differed in areas where the extensive margin (enrollment and composition) effects were likely to be largest. Our prior is that the effects of the policy on learning outcomes will be smaller, and perhaps even negative, the greater the evidence of gains in student access or of negative selection into the exam.

When examining aggregate outcomes, such as the quantity of test takers, we aggregate the data by district, as this is the relevant level for any public-private competition. These specifications also include district fixed effects because school effects are no longer identified.

We cluster all standard errors by region, the unit of treatment. Because there are only five regions in the sample, we conduct inference via the wild- t cluster bootstrap (Cameron et al. 2008), using the weights proposed by Webb (2013) for samples with fewer than 10 clusters. Results tables report p -values and significance levels based on these corrections. Because the bootstrap produces valid p -values but not confidence intervals (Cameron and Miller 2015, 27), we also report standard errors clustered by region based on the usual asymptotic approximation (Labonne 2013) but caution that these are illustrative and not appropriate for inference.

III. DATA

Sources

Outcome data are the universe of student exam records from the West African Examinations Council (WAEC). Subject-level scores are available for each student registered for the exam between 1998 and 2012, allowing for several years of pretreatment outcomes for each region.⁸ In addition to omitting region 1, we also omit private schools because they were ineligible for the scholarship program. However, all raw test results are converted to z -scores based on the universe of results in a given year, including students from private schools and region 1. This standardization allows us to

interpret scores relative to the national norm. It also explains why mean z -scores tend to be negative in our estimation sample.

Although our primary interest is the population of schools and students eligible for the program, nonrandom sorting of students into public and private schools in response to the scholarship may bias results. Such sorting is also an interesting potential outcome to investigate. In 2004, when program rollout was complete, only two of 43 districts had both a public and private high school (grades 10–12). By 2012, the last year for which we have data, this figure had grown to nine. We will therefore assess whether the growth in private school enrollment was related to the scholarship program.

However, all private schools are located in the urban districts of region 2, near the capital.⁹ Students in most areas are therefore constrained to attend their local public school. We later check robustness of results to various definitions of the estimation sample, such as excluding region 2 or including private schools.

We also use data from the 1998 wave of the Integrated Household Survey (IHS) to explore heterogeneity in results by baseline characteristics. This survey, which is conducted by the Gambia Bureau of Statistics, is nationally representative and collects information on assets, demographics, and socioeconomic characteristics. In the 1998 survey, slightly over 1,900 households were covered including approximately 4,500 school-aged children. A third and final data set we use in the analysis is the annual school census conducted by the Ministry of Education, which spans the same years as the exam data. The Ministry reports enrollment by grade and gender for each school, and the number and gender of teachers.

Unfortunately, we are unable to link individuals across these three data sets, preventing us from connecting individual enrollment decisions or household characteristics with test results. We therefore use district and school characteristics to analyze treatment effect heterogeneity. At the district level, we construct three measures using the 1998 IHS. We define “low enrollment” districts as those that fell below the national median enrollment rate for secondary school-aged children (ages 13–18).¹⁰ “Rural” districts are above the median percentage of population living in a rural area. “Wealthy” districts are above the median level of average household assets, with assets measured as the first principal component of ownership dummies for bicycle, car, refrigerator, motorcycle, sewing iron, television, radio, and VCR. At the school level, “distant” schools are located beyond the median distance from a main road. Importantly, all of these characteristics are predetermined with respect to the policy. All are also dummy variables, allowing for easy interpretation of interactions with the treatment indicator.

The nature of the scholarship program poses a potential challenge regarding data quality. Because the Ministry of Education remitted fees to schools for each girl enrolled, schools have an incentive to over-report enrollment in order to attract more resources (Sandefur and Glassman 2015).¹¹ Our reliance on exam records alleviates this concern, because the examination body WAEC is independent of the Ministry and requires separate student registration rather than automatically enrolling those on school rosters. The Ministry

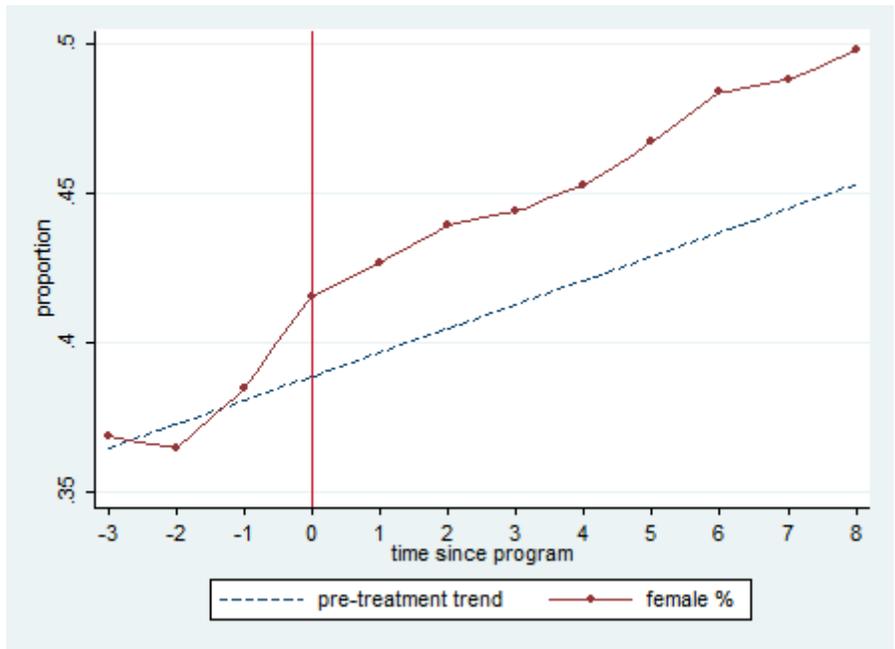
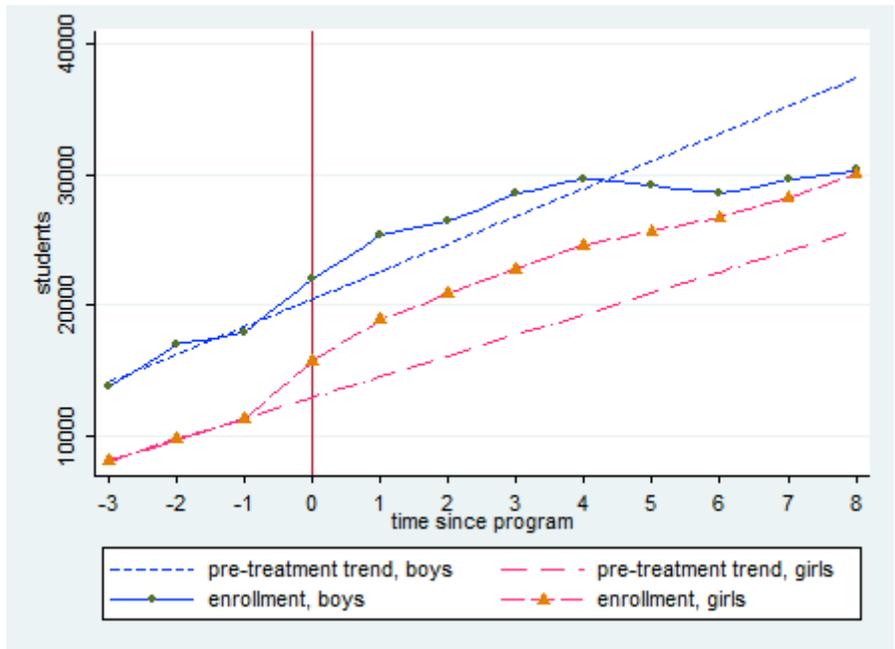
monitors its school enrollment records each year through site visits conducted by officials based in each region and in the central office, rather than relying on self-reports from school administrators. Nonetheless, in recognition of concerns about over-reporting we rely on enrollment data from the Ministry sparsely, largely for descriptive statistics. We address the potential bias from enrollment data used in more formal analysis when discussing results.

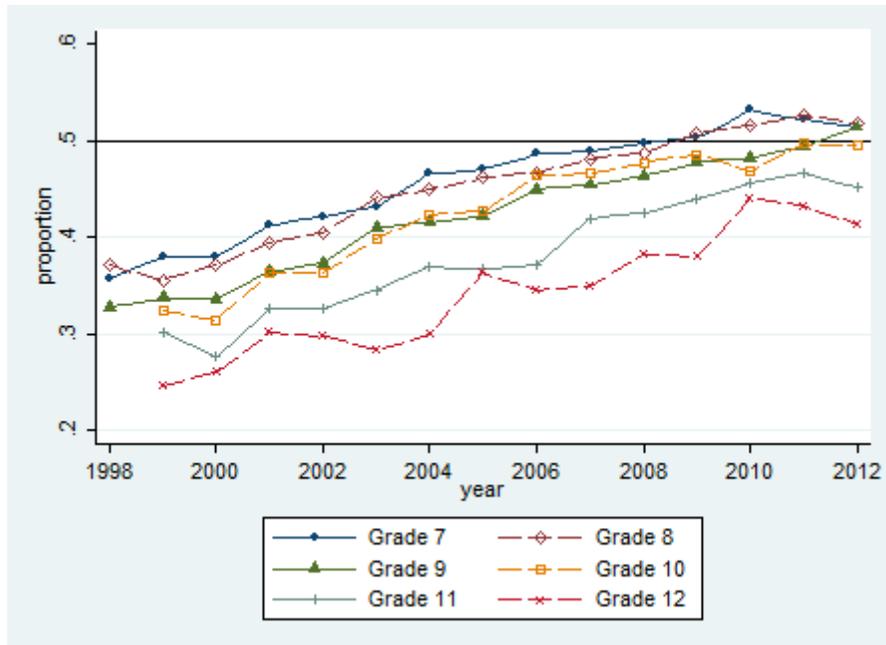
Descriptive Statistics

Given our focus on grade 12 outcomes in this paper, it is instructive to understand how Gambian students' progress through secondary school. The enrollment data do not track individual students over time, preventing us from constructing true grade progression rates. However, we can approximate progression through secondary school by comparing enrollment totals in grade 7, the first year of secondary school, with grade 10 enrollment three years later, when students transition from Upper Basic to Senior Secondary School. We can do the same for grade 12 enrollment and test taking five years later. Although these estimates will be biased due to mortality, migration, and grade repetition, they nonetheless give a sense of secondary school continuation and completion. For the seventh grade cohort entering in 2000, the year before the program began, enrollment in tenth grade three years later was only 29% of initial cohort size. Grade 12 enrollment and test-taking were each 25% of the initial seventh grade enrollment. These estimates suggest that secondary school progression is rare, but that students who persist to the upper secondary grades are likely to continue. Table S1 of the supplemental appendix presents these results, with additional breakdowns by gender, urban/rural, and region, as well as how grade progression changed over time for subsequent seventh grade cohorts.

The Gambia made considerable strides in reducing the gender enrollment gap since implementing the scholarship program. Figure 3 shows Ministry of Education data on enrollment in grades 7–12, aggregated across all public schools in regions 2–6. Panels (a)–(b) rescale time so that $t = 0$ corresponds to the first year of program receipt. Panel (a) shows that female enrollment increased relative to the pre-treatment trend after introduction of the program, while male enrollment fell.¹² Panel (b) shows the resulting increase in the female enrollment share. This is suggestive evidence of the program's effect on enrollment, consistent with Gajigo (2014), albeit subject to the potential bias in enrollment data discussed previously. Panel (c), which uses calendar time and disaggregates the data by grade, shows that the female enrollment percentage increased over time for all grades. It also shows that females comprise a lower share of enrollment as grade level increases (with only a few exceptions), meaning that females will be under-represented among test takers relative to their enrollment shares in their corresponding schools.

Figure 3(a)-(c): Secondary school enrollment, by gender (a) Enrollment, (b) Female enrollment proportion, (c) Female enrollment proportion, by grade. Figure shows secondary school (grade 7–12) enrollment. Data source: Gambia Ministry of Basic and Secondary Education.





Test-taking patterns follow the general upward enrollment trends of figure 3. Table 1 presents summary statistics separately for boys and girls at various points in time, for all sample regions and broken down by the region groups that received the policy. The number of test takers is relatively small in 1998, particularly for girls; the 210 girls taking the test that year represent only 22% of the total. By 2006, when all regions had the girls' scholarship program for at least three years, the number of girls taking the exam had nearly tripled and the female share rose to 36%. The growth in female test takers was particularly fast in regions 3–6, the earliest program regions and the most remote. These upward trends continued through 2010, though the growth in female test takers slowed. For both girls and boys, English and math scores improved over time, although the trend was not monotonic across all regions and years.

Table 1: Summary statistics, Grade 12 exam

	1998	2002	2006	2010
<u>Panel A: female test takers</u>				
total	210	382	624	820
region				
2	113	146	239	305
3-4	42	162	202	242
5-6	55	74	183	273
<u>Panel B: male test takers</u>				
total	749	846	1,130	1,371
region				
2	282	251	486	532
3-4	210	356	272	323
5-6	257	239	372	516
<u>Panel C: female share of test takers</u>				
total	0.22	0.31	0.36	0.37
region				
2	0.29	0.37	0.33	0.36
3-4	0.17	0.31	0.43	0.43
5-6	0.18	0.24	0.33	0.35
<u>Panel D: English and Math score, female</u>				
total	-0.46	-0.24	-0.28	-0.23
region				
2	-0.33	-0.01	-0.15	-0.04
3-4	-0.64	-0.44	-0.40	-0.36
5-6	-0.60	-0.27	-0.33	-0.32
<u>Panel E: English and Math score, male</u>				
total	-0.21	-0.14	-0.18	-0.09
region				
2	-0.01	0.11	-0.04	0.01
3-4	-0.51	-0.37	-0.30	-0.23
5-6	-0.18	-0.05	-0.26	-0.11

Sample is Grade 12 student exam records, public schools Regions 2-6, 1998-2012. Test scores are mean z -scores, standardized based on universe of students taking exam in that year, including Region 1.

IV. RESULTS

Pretreatment Outcome Trends

Before presenting estimates of the program's impact, we first check the validity of our identifying assumption of common outcome trends between regions that received the program early and those that received it late. Table 2 presents estimates of pretreatment trends from equation (2) for the quantity and performance of students taking the grade 12 exam. In addition to coefficients on the interactions between the time trend and region-groups (β_1 and β_2 from equation (2)), the bottom of the table presents p -values for tests of the joint hypothesis that these coefficients equal zero or that they are equal to each other; rejections of either of these hypotheses would be evidence of differential pre-treatment trends.

Table 2: Pre-treatment outcome trends

sample outcome units/subject	girls					boys				
	test takers		test scores			test takers		test scores		
	level	log	English & Math	English	Math	level	log	English & Math	English	Math
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
time to treatment*1(region=3,4)	7.2 (6.3) [.36]	0.29 (0.24) [.37]	0.01 (0.09) [.84]	-0.01 (0.10) [.80]	0.003 (0.05) [.97]	14.3 (15.7) [.37]	0.20 (0.17) [.37]	0.04 (0.04) [.69]	0.03 (0.05) [.73]	0.03 (0.01) [.52]
time to treatment*1(region=5,6)	4.1 (2.5) [.32]	0.11 (0.08) [.77]	0.03 (0.02) [.50]	0.04 (0.04) [.49]	-0.03* (0.01) [.07]	3.9 (11.6) [.91]	0.06 (0.13) [.88]	-0.04 (0.02) [.59]	-0.07 (0.03) [.17]	-0.02 (0.02) [.84]
N	37	37	1,380	1,380	1,380	37	37	3,738	3,738	3,738
R-squared	0.29	0.41	0.40	0.30	0.33	0.25	0.25	0.37	0.27	0.30
mean outcome	36.5	3.4	-0.27	-0.28	-0.19	96.4	4.5	-0.17	-0.22	-0.09
p-value on region-specific trends	[.97]	[.52]	[.55]	[.49]	[.98]	[.49]	[.57]	[.31]	[.94]	[.78]
p-value on Region 5-6 v. Region 3-4	[.85]	[.91]	[.35]	[.68]	[.82]	[.96]	[.87]	[.97]	[.96]	[.60]

Outcomes for Grade 12 exam shown. All regressions use pre-treatment years only, with time rescaled within region relative to first year of girls scholarship program. All regressions include linear trend in this rescaled time and calendar year fixed effects. Sample for level and log of test takers is district panel, and includes district fixed effects. Sample for test scores is students, and includes age, age squared, and school fixed effects. Standard errors in parentheses, clustered by region. p-values in brackets, calculated from 200 iterations of wild-t cluster bootstrap of Cameron et al (2008) with Webb (2013) 6-point distribution of weights. Significance levels are based on these p-values: * significant at 10%; ** significant at 5%; *** significant at 1%. p-value on region-specific trends refers to joint hypothesis test for significance of interactions between time trend and region dummies. p-value on Region 5-6 v. Region 3-4 refers to test of equality between coefficients on time trend-region interactions.

For the number of girls taking the exam (column 1), coefficients on the time trend-region interactions are not significant, either separately, jointly, or when comparing trends in regions 5–6 with regions 3–4. Using the log number of girls taking the exam as the outcome in column (2) to look for differential growth rates, we again find no evidence of pre-treatment trends. The analogous regressions for boys in columns (5)–(6) also provide no evidence of differential trends.

The remaining columns of the table present results for test scores at the student level, combining the English and math score and for each subject separately. Only one significant coefficient appears in the table (for girls’ math score in regions 5–6, at 10%), fewer than what we would expect to find by chance across the 20 coefficients tested in the table. We conclude that there is no evidence of differential pretreatment trends by region.¹³

Test Taking and Sorting in Response to Scholarship

The scholarship program could affect learning outcomes by altering the quantity of students taking exams, the composition of students, or the learning resources available to them. We examine each of these channels before presenting the main results.

In table 3, we present estimates of the effect of the girls’ scholarship program on the number of test takers. In panel A, column (1), the coefficient on program receipt indicates that in public (i.e., scholarship-eligible) schools, 26.1 more girls per district took the exam in regions that received the program early relative to those that did not, significant at 5%. Column (3) shows that this translated into approximately 55% more girls taking the exam in response to the program, also significant at 5%. The analogous increases for boys in column (2) and (4) were 42.6 students and 39%, though neither is statistically distinguishable from zero. The increase in the female share of all test takers (column 5) was also not statistically significant.¹⁴

Table 3: Effect of girls' scholarship program on number of test takers

outcome units sample	test takers						
	level		log		% female	as share of enrolled	
	girls (1)	boys (2)	girls (3)	boys (4)	(5)	girls (6)	boys (7)
Panel A: scholarship-eligible schools							
girls' scholarship program	26.1** (3.7) [.02]	42.6 (21.3) [.35]	0.55** (0.07) [.02]	0.39 (0.19) [.28]	0.03 (0.03) [.45]	0.28** (0.06) [.03]	0.13** (0.04) [.02]
R-squared	0.10	0.10	0.08	0.09	0.64	0.51	0.55
mean outcome	40.3	77.8	3.4	4.1	0.34	0.85	0.88
Panel B: all schools							
girls' scholarship program	96.1** (15.3) [.03]	112.9** (28.5) [.05]	0.93** (0.12) [.02]	0.69* (0.21) [.10]	0.02 (0.04) [.61]	0.23** (0.06) [.03]	0.16** (0.04) [.02]
R-squared	0.28	0.19	0.16	0.09	0.75	0.41	0.54
mean outcome	64.4	98.8	3.6	4.2	0.40	0.83	0.85
N	188	188	188	188	188	177	177

Sample is district panel, 1998-2012. All regressions include district and year fixed effects. Regressions for proportion female and share enrolled weighted by number of test takers. Number of observations differs for enrollment share due to districts with missing enrollment data. Standard errors in parentheses, clustered by region. *p*-values in brackets, calculated from 200 iterations of wild-*t* cluster bootstrap of Cameron et al (2008) with Webb (2013) 6-point distribution of weights. Significance levels are based on these *p*-values: * significant at 10%; ** significant at 5%; *** significant at 1%.

We lack direct measures of secondary school completion or university enrollment, because these are set by each institution and may vary over time. Nonetheless, increases in test takers suggest that the program brought additional students close to completion, complementing the findings of Gajigo (2014) on enrollment gains from the program. Another indication of completion is the share of students enrolled in grade 12 who attempt the exit exam. In panel A, columns (6)–(7), we find that the program led to a large increase in this share, 28 percentage points for girls and 13 percentage points for boys, both significant at 5%. Given the steep declines in enrollment (both absolute and female share) during the progression through secondary school, the persistence of students in response to the program is notable.¹⁵ These results are particularly striking because they run counter to the misreporting bias discussed in section III. If enrollment counts are over-reported in response to the policy but test registration is not (as we have reason to suspect), then the measured share of enrolled students taking the exam should fall. Additionally, if students induced to enroll by the policy are negatively selected and schools discourage weaker students from taking the exam, then the share of students taking the exam should fall further, but in fact we find large increases.¹⁶

Table 3, panel B repeats the regressions of panel A but for all schools in a district, as in the education market approach of Hsieh and Urquiola (2006). The increases in test-taking found in panel A are magnified when all schools are included, such that the program effect increases to 96.1 additional girls taking the test (column 1), or an increase of approximately 93% (column 3) for girls, both significant at 5%. These results suggest that the program induced an exodus of students to private schools, as in the study of free primary schooling in Kenya by Lucas and Mbiti (2012a). Additionally, we find positive and significant effects for boys' test taking when considering all schools in column (2) and (4), with increases of 112.9 in levels and 69% in proportions. We will explore the mechanisms behind the spillover effect for boys later in the paper.

Table 4 explores the private school response further. The share of students taking the exam in private schools increased by 17 percentage points for girls and 21 percentage

points for boys, as shown in columns (1) and (5). However, columns (2) and (6) show that these increases were driven entirely by region 2, which is the most urban and wealthy of treated regions. The interaction effect between the program and a dummy for region 2 is positive and significant, as is the sum of the main effect of the program and this interaction (reported as “treatment effect with interaction” at the bottom of the table). The private school share for girls fell by 23 percentage points in regions 3–6, showing that the scholarship attracted students to public schools in these poorer regions. Columns (3) and (7) reveal that there was also no significant exit to private schools in districts with low enrollment. Private school increases were also concentrated in wealthy districts (treatment effect with interaction results for columns 4 and 8). Overall, then, the exit of students to private schools occurred only in the relatively advantaged areas where a contested education market existed.^{17,18}

Table 4: Private school response to girls’ scholarship program

outcome sample	private school share of test takers								number of private schools			
	girls				boys				(9)	(10)	(11)	(12)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)				
girls’ scholarship program	0.17*** (0.01) [.01]	-0.23** (0.04) [.05]	0.20** (0.04) [.02]	-0.14 (0.16) [.83]	0.23** (0.02) [.03]	-0.08 (0.02) [.17]	0.25** (0.04) [.03]	-0.02 (0.16) [.97]	-0.06 (0.19) [.73]	0.09 (0.14) [.60]	-0.07 (0.22) [.71]	0.27 (0.28) [.44]
interaction of program with: Region 2		0.56*** (0.02) [.00]				0.47*** (0.004) [.00]				-0.25 (0.12) [.27]		
low enrollment			-0.36 (0.18) [.29]				-0.26 (0.15) [.24]				0.06 (0.22) [.82]	
wealthy				0.36 (0.19) [.50]				0.29 (0.19) [.49]				-0.45 (0.19) [.32]
N	184	184	184	184	184	184	184	184	184	184	184	184
R-squared	0.86	0.90	0.87	0.88	0.83	0.91	0.84	0.85	0.34	1.00	1.00	1.00
treatment effect with interaction		0.32** (0.18) [.03]	-0.15 (0.14) [.40]	0.22*** (0.14) [.01]		0.40*** (0.24) [.01]	-0.004 (0.07) [.90]	0.27*** (0.17) [.01]		-0.16 (0.24) [.54]	-0.01 (0.17) [.84]	-0.18 (0.21) [.43]
mean outcome	0.38	0.38	0.38	0.38	0.22	0.22	0.22	0.22	12.30	12.30	12.30	12.30

Outcome is share of students taking Grade 12 exam in private schools. Sample is district panel, Regions 2-6, 1998-2012. Schools is number of private secondary schools. Low enrollment and wealth interactions are district characteristics from 1998 household survey. Low enrollment = below national median secondary-age (13-18) enrollment rate. Wealthy = at or above national median household assets, by first principal component. All regressions include district and year fixed effects. “Treatment effect with interaction” is sum of program coefficient and coefficient on interaction term. Standard errors in parentheses, clustered by region. *p*-values in brackets, calculated from 200 iterations of wild-*t* cluster bootstrap of Cameron et al (2008) with Webb (2013) 6-point distribution of weights. Significance levels are based on these *p*-values: * significant at 10%; ** significant at 5%; *** significant at 1%.

Given these increases in the number of test takers, the share of students enrolled in grade 12 taking the exam, and their shift to private schools, we are also interested in whether the characteristics of test takers changed. A useful proxy for the quality of a student taking the exam is age. Students who are old for their grade are more likely to have started school late, repeated grades, or had their schooling interrupted by periods of nonenrollment, all of which are likely to indicate negative selection into test-taking relative to the average student. Consistent with this hypothesis, in pretreatment periods test scores declined steadily with age, from a mean *z*-score of 0.41 for girls aged 16 to a mean score of -0.60 at age 24.¹⁹ If the age distribution of test takers is systematically related to treatment, this would indicate that the policy altered the composition of students selecting into the exam.

Table 5 explores whether the scholarship program changed the composition of test takers within public schools. Using individual student exam records, we define the outcome as a dummy for whether the test taker is more than 20 years old, an age threshold that roughly

marks whether the student takes the exam “on time” based on typical school progression. In column (1), we find that among girls taking the exam, the program increased the share of students older than 20 by 5 percentage points, though the coefficient is not statistically significant. For boys, the point estimate is similar but also not significant (column 2). A triple-difference specification that pools all students and compares girls and boys in program regions also produces no significant effects of the program (column 3).

Table 5: Effect of girls' scholarship program on age of test takers

outcome sample	age > 20										
	girls (1)	boys (2)	all (3)	girls				boys			
	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)			
girls' scholarship program	0.05 (0.02) [.26]	0.04 (0.05) [.53]	0.05 (0.03) [.29]	0.04 (0.03) [.31]	0.02 (0.03) [.42]	-0.01 (0.04) [.92]	0.13** (0.02) [.04]	0.06 (0.03) [.31]	0.06** (0.03) [.05]	0.01 (0.04) [.86]	0.17*** (0.02) [.01]
interaction of program with:											
male			-0.02 (0.01) [.20]								
low enrollment				0.11 (0.05) [.20]				0.13 (0.02) [.11]			
rural					0.09 (0.04) [.86]				0.06 (0.02) [.63]		
distant						0.10 (0.06) [.40]				0.06 (0.05) [.45]	
wealthy							-0.10 (0.02) [.29]				-0.11 (0.02) [.14]
N	7,997	15,811	23,808	7,234	7,234	7,997	7,234	13,813	13,813	15,811	13,813
R-squared	0.04	0.03	0.05	0.03	0.03	0.04	0.03	0.03	0.03	0.03	0.03
treatment effect			0.04	0.14**	0.11	0.10**	0.03	0.19**	0.12	0.06	0.06
with interaction			(0.03) [.40]	(0.04) [.05]	(0.04) [.37]	(0.02) [.02]	(0.03) [.31]	(0.02) [.02]	(0.04) [.11]	(0.07) [.46]	(0.04) [.29]
mean outcome	0.35	0.51	0.45	0.36	0.36	0.35	0.36	0.52	0.52	0.51	0.52

Outcome is dummy for student age above 20. Sample is students taking grade 12 exit exam in Regions 2-6, 1998-2012. Interaction with male based on individual student. Distant is dummy for above national median distance from school to main road. All other interactions are district characteristics from 1998 household survey. Low enrollment = below national median secondary-age (13-18) enrollment rate. Rural is above national median proportion of 1998 district population living in rural area. Wealthy = at or above national median household assets, by first principal component. “Treatment effect with interaction” is sum of program coefficient and coefficient on interaction term. All regressions include school and year fixed effects. Regression with both male and female students in sample also includes female dummy. Sample sizes differ across regressions because not all districts surveyed in 1998. Standard errors in parentheses, clustered by region. *p*-values in brackets, calculated from 200 iterations of wild-*t* cluster bootstrap of Cameron et al (2008) with Webb (2013) 6-point distribution of weights. Starred significance levels are based on these *p*-values. * significant at 10%; ** significant at 5%; *** significant at 1%.

We look for heterogeneity in the age distribution according to district characteristics in the remainder of the table. In column (4), we find that girls in low enrollment districts are 14 percentage points more likely to be older than 20 in response to the policy, significant at 5% (“treatment effect with interaction” reported at bottom of table). Girls in distant schools are also 10 percentage points more likely to be older than 20 in response to the policy (column 6). In the specification with an interaction of treatment with wealthy districts in column (7), the coefficient on the main effect of the program means that girls taking the exam in poorer districts are 13 percentage points more likely to be older than 20, significant at 5%. The positive interaction term with rural in column (5) is also consistent with negative selection into the test, though it is not statistically significant. Point estimates for boys follow the same pattern, with large and significant increases in older students in low enrollment, more urban, and poorer districts (19, 6, and 17 percentage points, respectively). In sum, the results in table 5 show that the program increased the proportion of older test takers in more disadvantaged areas, consistent with negative selection.

In addition to changes in the number and composition of students, changes in enrollment can also be accompanied by changes in school quality. For instance, increases in pupil-

teacher ratio would indicate if the influx of test takers due to the program strained teaching capacity. Changes in the proportion of teachers who are female would reveal whether the program's focus on female students also influenced the gender composition of teachers, an important element of school quality for girls (Muralidharan and Sheth 2015). We find no significant effects of the program on these school quality measures, either overall or by district characteristics. See table S4 of the supplemental appendix.

Student Learning

As discussed earlier, the effect of the girls' scholarship program on student learning is theoretically ambiguous. Fee elimination could reduce stress or free students from the need to engage in income-generating activity, thereby improving performance. On the other hand, an influx of new students could lower the average quality of students or place strain on school resources and harm the learning environment. We documented several changes to the quantity and composition of students in response to the program in tables 3–5. First, the number of girls taking the test increased. Second, the most developed and urban region of the sample (region 2) saw an increase in the market share of private schools. Third, the share of older students taking the exam increased in more disadvantaged areas. All else equal, each of these trends suggests average student performance should fall.

Table 6 presents results from the main difference-in-difference regression (1), using a student's English and math score as the outcome and the same specifications as table 5. The point estimates in columns (1)–(2) show that the program increased scores by .09 standard deviations for girls and .11 standard deviations for boys. Although neither effect is statistically different from zero, the absence of statistically significant negative effects of the policy on learning outcomes is notable, given our earlier findings on student enrollment and composition. The triple difference specification in column (3) also reveals no significant program effects.

Table 6: Effect of girls' scholarship program on test scores

outcome sample	English & Math										
	girls (1)	boys (2)	all (3)	girls				boys			
	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)			
girls' scholarship program	0.09 (0.03) [.11]	0.11 (0.06) [.37]	0.12 (0.06) [.35]	0.05 (0.02) [.33]	0.13* (0.03) [.10]	0.18* (0.05) [.06]	-0.05 (0.08) [.86]	0.05 (0.04) [.45]	0.12* (0.03) [.10]	0.10 (0.06) [.65]	-0.01 (0.10) [.88]
interaction of program with:											
male			-0.02 (0.02) [.42]								
low enrollment				0.04 (0.03) [.35]				0.05 (0.06) [.75]			
rural					-0.23 (0.03) [.14]				-0.22* (0.02) [.07]		
distant						-0.16 (0.09) [.45]				0.03 (0.17) [.75]	
wealthy							0.13 (0.10) [.83]				0.09 (0.14) [.64]
N	7,997	15,811	23,808	7,234	7,234	7,997	7,234	13,813	13,813	15,811	13,813
R-squared	0.35	0.27	0.29	0.37	0.37	0.35	0.37	0.31	0.31	0.27	0.31
treatment effect			0.10	0.09**	-0.10	0.02	0.09	0.10	-0.10	0.13	0.07
with interaction			(0.05) [.34]	(0.03) [.02]	(0.02) [.52]	(0.06) [.83]	(0.03) [.30]	(0.06) [.31]	(0.04) [.32]	(0.13) [.69]	(0.05) [.54]
mean outcome	-0.28	-0.14	-0.19	-0.27	-0.27	-0.28	-0.27	-0.16	-0.16	-0.14	-0.16

Outcome is Grade 12 exam score in English and Math, standardized by universe of scores in that year. Sample is students taking grade 12 exit exam in Regions 2-6, 1998-2012. Interaction with female based on individual student. Distant is dummy for above national median distance from school to main road. All other interactions are district characteristics from 1998 household survey. Low enrollment = below national median secondary-age (13-18) enrollment rate. Rural is above national median proportion of 1998 district population living in rural area. Wealthy = at or national above median household assets, by first principal component. "Treatment effect with interaction" is sum of program coefficient and coefficient on interaction term. All regressions include age, age squared, school and year fixed effects. Regression with both male and female students in sample also includes female dummy. Sample sizes differ across regressions because not all districts surveyed in 1998. Standard errors in parentheses, clustered by region. *p*-values in brackets, calculated from 200 iterations of wild-*t* cluster bootstrap of Cameron et al (2008) with Webb (2013) 6-point distribution of weights. Starred significance levels are based on these *p*-values. * significant at 10%; ** significant at 5%; *** significant at 1%.

Column (4) reveals that in low enrollment districts, the program increased girls' test scores by .09 standard deviations ("treatment effect with interaction" reported at bottom of table, significant at 5%). This increase is particularly notable given that the share of older girls taking the exam in these districts, an indicator of negative selection, increased by 14 percentage points. A potential explanation is that these older girls were in fact stronger students who had to leave school for financial reasons, but re-entered due to the policy. In column (5), the main effect of the program indicates an increase of .13 standard deviations for girls in more urban areas, with no significant difference in scores in treated rural areas. Column (6) shows a similar pattern of results by school distance, with scores at less distant schools increasing .18 standard deviations in response to the program. These increases are consistent with the intensive margin response—that is, alleviating financial stress among students who would have taken the exam in the absence of the program—dominating the extensive margin response in more urban schools.

We find no significant treatment effects according to district wealth in column (7). As with the results by low enrollment, the absence of significant test score declines in poorer districts and more distant schools is notable given the corresponding increase in older girls taking the exam in those areas. The apparent negative selection into the exam found in table 5 did not translate into test score declines and in fact did not prevent modest test score increases in some areas.

Results for boys in columns (8)–(11) show no significant effects on test scores, with the exception of a .12 standard deviation increase in more urban districts, significant at 10%. The positive coefficient is noteworthy given the 6 percentage point increase in the share of older students in these districts. The negative interaction term for rural fails to produce

a statistically significant overall effect for rural districts; see bottom of table. The absence of test score declines for boys is arguably even more striking than for girls, given the more pronounced pattern of negative selection according to age found for boys in table 5.²⁰

A potential explanation for our results is that students may have reduced their effort on other exam subjects in order to focus on the required English and math sections. However, we find no decrease in the total number of exam subjects taken by students in response to the program, for either girls or boys. Nor do we find changes in the share of “easy” subjects taken in response to the program, where “easy” corresponds to a subject with a passing rate above the median in the pretreatment data. These results, reported in table S5 of the supplemental appendix, reveal no evidence of decreased student effort.

Spillover Effects on Boys

Among the results presented thus far, the evidence of increased test taking for boys, with no corresponding decrease in achievement, in response to a program that targeted girls is perhaps the most intriguing. What explains this spillover effect? One possibility is that the scholarship alters constraints on human capital investment within a household. The opportunity to send a girl to school without incurring fees frees resources to send boys to school. Alternately, parental preferences for equivalent treatment of children, or a desire for boys to accompany their sisters, could induce a similar response.

In this subsection, we look for evidence of such intra-household spillovers in response to the program. We distinguish whether a secondary school-aged boy lives with younger or older girls, as we expect that the causes of spillovers might differ in the two cases. Spillovers from older girls to younger boys are more likely to reflect financial considerations, as liquidity-constrained households with older girls would gain relief earlier. We expect spillovers for nonfinancial reasons to be stronger when girls are younger, as parents would prefer older boys to accompany younger girls to school.

Unfortunately, the exam records do not identify which students belong to the same household, preventing us from testing these hypotheses using the same outcomes already considered. However, the Gambian Integrated Household Survey data allow us to explore the enrollment response using a richer set of individual and household characteristics. We use the 1992, 1998, 2003, and 2010 waves of the survey, limiting the sample to secondary school-aged boys (13–18) in regions 2–6. The data record all household members, allowing us to observe if boys live in the same household as scholarship-eligible girls (though we cannot distinguish siblings from other types of connections).

We run a series of linear probability models in which the outcome is an indicator for enrollment. Note that this outcome differs from others previously considered, because the sample is all boys in the age group, whereas previously our sample was Grade 12 boys taking the exit exam. In table 7, column (1), we find no statistically significant effect of the program on boys’ enrollment. The finding is consistent with Gajigo (2014), who found no increase in the boys’ enrollment rate in response to the program.

Table 7: Boys' enrollment spillover effects, by girls' age

	boys' enrollment		
	(1)	(2)	(3)
girls' scholarship program	0.02 (0.03) [.54]	0.01 (0.04) [.71]	0.02 (0.03) [1.00]
older girl in HH of secondary school age		-0.01** (0.01) [.03]	
younger girl in HH of secondary school age		-0.01 (0.02) [1.00]	
older girl in HH enrolled in secondary school			0.14 (0.04) [.13]
younger girl in HH enrolled in secondary school			0.10 (0.05) [.29]
older girl in HH of secondary school age*program		0.03 (0.03) [.41]	
younger girl in HH of secondary school age*program		-0.004 (0.04) [.99]	
older girl in HH enrolled in secondary school*program			0.07 (0.02) [.14]
younger girl in HH enrolled in secondary school*program			0.08 (0.06) [.34]
<i>N</i>	4,718	4,718	4,718
R-squared	0.12	0.12	0.14
program spillover effect (older girl in HH)		0.04 (0.04) [.64]	0.10*** (0.04) [.00]
program spillover effect (younger girl in HH)		0.02 (0.03) [.42]	0.10 (0.06) [.12]
mean outcome	0.60	0.60	0.60

Sample is boys of secondary school age (13-18), Regions 2-6, Gambia Integrated Household Survey 1992/1998/2003/2010. "Older" girls include those of same age as boy. Program spillover effect is sum of program and program*girl in HH coefficients. Standard errors in parenthesis, clustered by region. *p*-values in brackets, calculated from 200 iterations of wild-*t* cluster bootstrap of Cameron et al (2008) with Webb (2013) 6-point distribution of weights. Starred significance levels are based on these *p*-values. * significant at 10%; ** significant at 5%; *** significant at 1%.

Spillovers might occur if the presence of a scholarship-eligible girl in the household changes boys' enrollment. In column (2), there is a modest but statistically significant decrease (1 percentage point) in boys' enrollment in households with an older girl of secondary school age, consistent with sibling rivalry in which the older children are the first recipients of household education expenditure. This decline in boys' enrollment is

not present in treated areas, however, as the sum of the coefficients on the program coefficient and its interaction with the dummy for older girl is not statistically significant (“program spillover effect (older girl in HH)” at bottom of table). The discrepancy in these results is consistent with the scholarship alleviating financial constraints.

In column (3), we find that the program increased enrollment by 10 percentage points for boys in households with an older girl enrolled in secondary school, not merely scholarship-eligible.²¹ This spillover effect cannot be explained by unobserved differences in household preferences for schooling, which is captured by the included main effects of girls’ secondary enrollment (which are also positive, though not significant). Instead, comparing columns (2) and (3) shows that scholarship-eligible households with older girls that take up the program also increase boys’ enrollment. Although girls’ enrollment is endogenous, the differential magnitude of the program effect among households with enrolled girls is consistent with the scholarship alleviating financial constraints.

Robustness Checks

Table 8 presents a series of robustness checks of the main test score results using alternative definitions of the estimation sample, with panel A for girls and panel B for boys. In column (1), we restrict the sample to regions 3–6, given the concentration of private schools in region 2.²² Point estimates are larger than those in table 6 for both girls and boys, with the .11 standard deviation increase for girls statistically significant at 1%. Column (2) pools all regions, including region 1, which never received the program. Point estimates are positive but not significant.

Table 8: Effect of girls' scholarship on test scores, alternative samples

outcome sample	English & Math					
	<u>exclude</u>	<u>all</u>	<u>include</u>	<u>all regions and</u>	<u>age ≤ 20</u>	<u>age > 20</u>
	Region 2	regions	private schools	private schools		
	(1)	(2)	(3)	(4)	(5)	(6)
<u>Panel A: girls</u>						
program receipt	0.11*** (0.03) [.01]	0.14 (0.03) [.11]	0.06 (0.03) [.22]	0.08* (0.02) [.07]	0.09 (0.03) [.20]	0.07* (0.02) [.07]
Observations	4,836	20,866	12,337	31,645	5,155	2,842
R-squared	0.10	0.29	0.32	0.42	0.38	0.25
<u>Panel B: boys</u>						
program receipt	0.18 (0.08) [.23]	0.04 (0.05) [.25]	0.11 (0.06) [.34]	0.02 (0.05) [.77]	0.08 (0.10) [.84]	0.13** (0.02) [.02]
N	9,938	34,872	19,509	43,791	7,704	8,107
R-squared	0.10	0.25	0.26	0.32	0.28	0.20

Table shows Grade 12 student test scores, standardized by universe of scores within year, regressed on girls' scholarship program, age, age squared, and school and year fixed effects. Columns (3), (5) and (6) include Regions 2-6 only. "On-time" students in column (5) are under age 20, "older" students in column (6) are 20 or older. Standard errors in parentheses, clustered by region. *p*-values in brackets, calculated from 200 iterations of wild-*t* cluster bootstrap of Cameron et al (2008) with Webb (2013) 6-point distribution of weights. Starred significance levels are based on these *p*-values. * significant at 10%; ** significant at 5%; *** significant at 1%.

Column (3) restricts the sample to regions 2–6 but includes students in private schools. These regressions continue to define treatment at the regional level, so that female private school students in treated regions are considered program recipients even though they must pay school fees, in order to mitigate confounding variation due to non-random sorting into private schools, as in the education market approach of Hsieh and Urquiola (2006). Once again, point estimates are positive but not distinguishable from zero. Column (4) includes all regions and private schools, with the coefficient for girls of nearly identical magnitude as in table 6 (.08 vs. .09 in the original specification) and significant at 10%.

Column (5) limits to the sample to students younger than 20, a group more likely to attend school in the absence of the program. Coefficients are similar in magnitude to the full sample but not significant. In column (6) we restrict the sample to students older than 20, a group more likely to be negatively selected, and which we showed in table 5 increased its presence due to the policy. The program coefficient is .07 for girls and .13 for boys, significant at 10% and 5%, respectively. Together, columns (5)–(6) suggest that changes in student composition induced by the program did not reduce performance at different points in the age distribution, and may even have increased performance for older students. An alternative explanation is that older students induced into the exam due to the policy were in fact positively selected, which could be the case if stronger students whose studies were interrupted by financial constraints re-entered due to the policy.

Failure to find any significantly negative effects of the policy for any gender or estimation sample considered in table 8 is notable, given the potential channels through

which the policy could reduce average test scores. These findings increase our confidence that the learning gains found for various subgroups in the preferred specifications of table 6 are not masking learning declines among other major subgroups of students.²³

V. CONCLUSION

This paper evaluated the effect of the Gambian girls' scholarship program on the quantity, composition, and achievement of secondary school students. Our approach relied on difference-in-differences estimation, comparing regions that received the program early to those that received it late. We validated this identification strategy by verifying that outcome trends were similar across regions prior to treatment. We found that the number of girls taking the high school exit exam increased due to the girls' scholarship program, consistent with the presence of financial constraints on enrollment in secondary school. Our results complement those of Gajigo (2014), who found increased enrollment among girls aged 13–18 and extend them in two important ways. First, because our results are based on the number of students sitting the grade 12 exit exam, they demonstrate that the effects of the scholarship program persisted throughout secondary school, rather than being limited to earlier grades. Second, we find evidence of increased access for boys as well, with gains in test taking among older boys, a group that would have started school late, repeated grades, or had their studies interrupted.

We also find changes in the composition of students in response to the scholarship. The share of older students taking the exam increased in poorer districts, consistent with negative selection. Enrollment spillovers to boys occurred only in households with older girls enrolled in school, consistent with programs in other countries in which reduced schooling costs for girls increased male enrollment within a household (Kim et al. 1999a; Begum et al. 2012).

We find robustly positive point estimates of the policy effect on test scores for both genders and across many samples and specifications. The failure to find any negative effects on learning is striking given the expanded access to secondary school from the program. Some subgroups likely to be negatively selected—such as girls in low enrollment districts, and older students—experienced modest but statistically significant test score gains. Our interpretation is that any negative selection induced by fee elimination was not sufficient to reduce learning on average.

Our results suggest that improving access to secondary education in countries where enrollment is low need not come at the expense of student learning. As developing countries increasingly turn their attention to secondary school, finding policies to promote both opportunity and achievement should sit high on the agenda.

Footnotes

¹ Blimpo (2014) evaluated the effect of financial incentives on secondary school children in Benin and found large gains on test scores. This policy, however, did not target access directly and provided no additional resources upfront.

² Schools serve as test centers. In almost all cases, students take the exams at the school they attend.

³ The major sub-national units in The Gambia are 6 regions. Region 1 includes the capital Banjul, with regions 2–6 at increasing remove heading east along the Gambia River bisecting the country. Below these subnational units, there are 43 districts as of 2013.

⁴ For purposes of this paper, public schools refer to both government and grant-aided schools, the latter of which are publicly funded but administered privately. Both types of public schools are eligible for the scholarship program, while private schools are not.

⁵ We follow the Gambian convention in referring to the 2000–2001 academic year as 2001, to 2001–2002 as 2002, and so on.

⁶ The average value changed over time because of changes in the exchange rate (the average value of a US dollar per Gambian Dalasi was approximately 13, 15, 20, 27 between 2000 and 2003) and also changes in the composition of students covered (middle and high school students) over time as the program got scaled.

⁷ Although it is possible that girls were previously paying less than 100% of the nominal fee, leading to an increase in school resources due to the policy, denial of school services for non-payment was a common practice.

⁸ We omit exam data from 2004 because student gender is missing for that year.

⁹ The map in figure S1 of the supplemental appendix shows secondary schools in 2011, the most recent year for which location data are available. Region 1 schools and Upper Basic (middle) schools are excluded from the estimation sample but shown on the map to illustrate the locations of secondary schools throughout the system. Not all schools in the sample appear on the map due to missing location data.

¹⁰ The strategy resembles that of Lucas and Mbiti (2012a, 2012b), who study the effect of free primary schooling in Kenya using variation in pretreatment local enrollment rates.

¹¹ An alternative possibility mentioned by a referee is that the policy led to greater scrutiny of “ghost” students, reducing reported enrollment. The direction of misreporting in response to the policy is therefore unclear.

¹² Increased enrollment in private schools in later periods partially explains the fall in male enrollment. We explore private school enrollment later in the paper.

¹³ There is also no visual evidence of differential pretreatment trends across these region groupings in the raw data. See figure S2 of the supplemental appendix for plots of the mean test score and number of test takers by time to treatment.

¹⁴ We also looked for district-level heterogeneity in the effect of the program on test taking by interacting the program with dummies for low enrollment, rural, and wealthy districts in separate specifications. None of the interaction terms are statistically significant for girls or boys, indicating that increased access for girls was widely shared across districts. Results appear in supplemental appendix table S2.

¹⁵ We have also used administrative data from the Ministry of Education to explore the enrollment effect of the policy by grade. Not surprisingly, the effects are largest in the early secondary grades and diminish as the students' progress through school. Results appear in figure S3 of the supplementary appendix.

¹⁶ Another potential explanation for the results is that additional students retook the exam in response to the scholarship. However, the scholarship covered exam fees only for enrolled students, which would mute this response.

¹⁷ In columns (9)–(11), we also find no change in the number of private schools in response to the program, either overall or differentially by the characteristics considered. This suggests that private schools experiencing enrollment gains grew in size as a result of the program.

¹⁸ A potential explanation for the increases in test takers and private school share is migration in response to the program. Although Gambian household surveys lack data on residential migration that would allow us to test directly, we think that such a response is unlikely, because the monetary and psychic costs of migration should exceed the scholarship's value of less than US\$50 annually. Switching schools without changing residence is also unlikely given the sparse geographic distribution of schools, particularly in regions 3–6, as shown in supplemental appendix figure S1. If such switching occurs, it would most likely be at schools close to the border between an eligible and ineligible region. To test this possibility, we analyzed the number of test takers and share in private schools by district as in tables 3 and 4 but included an interaction effect between program receipt and whether the district was located on the border of a regional grouping with different program rollout dates (i.e., the borders between Regions 1/2, Regions 2/4, or Regions 3–4/5). If students switched schools to benefit from the program, then these border districts should see a differential change in enrollment. None of the interaction terms are significant, suggesting that such switching was not a common response. See supplemental appendix table S3 for results.

¹⁹ Scores for boys are similar; see figure S4 of the supplemental appendix. Regressing test scores on a full set of age, school, and year dummies in pre-treatment periods yields a similar pattern, although the age coefficients are noisy.

²⁰ We also ran the specifications in table 6 separately for English and math scores. Test scores' gains are concentrated in English, while declines in math are present for some groups (girls in rural and poorer districts). These differences may have arisen because math skills depreciate more rapidly than English among students who return to school in response to the program. Results appear in tables SA6–SA7 of the supplemental appendix.

²¹ The magnitude of the effect is similar to the finding in Gajigo (2014) of an 11-percentage point increase in girls' enrollment due to the program, suggesting that households prefer boys and girls to attend school together.

²² Another potential confounding factor in region 2 was the Ambassador Girls Scholarship Program, funded by USAID, which targeted students in the cohort entering grade 7 in 2007 at a subset of secondary schools in that region (Giordano and Pugatch 2015). It went beyond the program studied in this paper by covering books, uniforms, and school supplies. However, its recipients entered twelfth grade in 2012, the last year of our data, meaning that the two scholarship programs overlap for only one region-year in our sample, making it unlikely to alter this study's results.

²³ In an additional set of robustness checks, we analyze whether time since program implementation matters by running an event study specification, in which leads and lags of the treatment indicator allow the treatment effect to vary by years before or since treatment began. We define $t = 0$ to be the first year of treatment and include all region-years from 3 years before treatment to 8 years after, which are the periods of overlap among all regions according to this time scale. We set $t = -1$ as the omitted category. The results, presented in figure S5 of the supplemental appendix, show that point estimates for the number of students taking the exam and test scores change from negative to positive at the year of treatment, for both girls and boys. These effects fluctuate over time but remain positive in all post-treatment years. These results increase our confidence that the overall treatment effect estimated in the paper is robust to a more granular specification based on program timing.

REFERENCES

Baird, S., C. McIntosh, and B. Ozler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *Quarterly Journal of Economics* 126 (4): 1709–1753.

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* 122 (3), 1235–64.

Banerjee, A. V., P. Glewwe, S. Powers, and M. Wasserman. 2013. "Expanding Access and Increasing Student Learning in Post-Primary Education in Developing Countries: A Review of the Evidence." Abdul Latif Jameel Poverty Action Lab (USA). Published by J-PAL.

Begum, L., A. Islam, R. Smyth. 2012. "Girls' education, Stipend Programs and the Effects on Younger Siblings' Education." IZA Working Paper.

Blimpo, M. P., and D. K. Evans. 2011. "School-Based Management and Educational Outcomes: Lessons from a Randomized Field Experiment." enGender Impact : The World Bank's Gender Impact Evaluation Database. Washington, DC, World Bank.

Blimpo, M. 2014. "Team Incentives for Education in Developing Countries: A Randomized Field Experiment in Benin." *American Economic Journal: Applied Economics* 6 (4): 90–109.

Burde, D., and L. L. Linden. 2013. "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools." *American Economic Journal: Applied Economics* 5 (3): 2740.

Cameron, A. C., D. L. Miller. 2015. "A Practitioners Guide to Cluster-Robust Inference." *Journal of Human Resources* 50 (2): 317–72.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90 (3): 414–27.

Chaudhury, N., and D. Parajuli. 2010. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Programme on Public School Enrolments in Punjab, Pakistan." *Applied Economics* 42 (28–30): 3565–83.

Cullen, J. B., R. Reback. 2006. *Tinkering Toward Accolades: School Gaming Under a Performance Accountability System*. Working Paper 12286. National Bureau of Economic Research.

Daly, A., B. Mbenga, and A. Camara. 2014. "Barriers to Participation and Retention: Engaging and Returning out of School Children in the Gambia." *Education 3-13*, June, 1–16.

Duflo, E., P. Dupas, and M. Kremer. 2009. "Returns to Secondary Schooling in Ghana." URL: <http://www.povertyactionlab.org/evaluation/returns-secondary-schooling-ghana>.

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

Gajigo, O. 2014. "Closing the Education Gender Gap: Estimating the Impact of Girls' Scholarship Program in The Gambia." *Education Economics*, 1–22.

Giordono, L., and T. Pugatch. 2015. Informal Fee Elimination and Student Performance: Evidence from The Gambia.

Glewwe, Paul, and Muralidharan, Karthik. 2015. Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications.

Hsieh, Chang-Tai, and Urquiola, Miguel. 2006. The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of public Economics*, 90(8): 1477–1503.

Kazianga, Harounan, Levy, Dan, Linden, Leigh L., and Sloan, Matt. 2013. The Effects of 'Girl-Friendly' Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso. *American Economic Journal: Applied Economics*, 5(3): 41–62.

Kim, Jooseop, Alderman, Harold, and Orazem, Peter F. 1999a. Can Private School Subsidies Increase Enrollment for the Poor? The Quetta Urban Fellowship Program. *World Bank Economic Review*, 13(3): 443–465.

Kim, Jooseop, Alderman, Harold, and Orazem, Peter F. 1999b. Evaluation of the Balochistan Rural Girls' Fellowship Program-Will rural families pay to send girls to school?

Kremer, Michael, Miguel, Edward, and Thornton, Rebecca. 2009. INCENTIVES TO LEARN. *The Review of Economics and Statistics*, 91(3): 437–456.

Krishnaratne, S., White, H., and Carpenter, E. 2013. *Quality education for all children? What works in education in developing countries*. Tech. rept. 3ie Working Paper 20.

Labonne, Julien. 2013. The local electoral impacts of conditional cash transfers: Evidence from a field experiment. *Journal of Development Economics*, 104, 73–88.

Lucas, A. M., and I. M. Mbiti. 2012a. “Access, Sorting, and Achievement: The Short-Run Effects of Free Primary Education in Kenya.” *American Economic Journal: Applied Economics* 4 (4): 226–53.

———. 2012b. Does Free Primary Education Narrow Gender Differences in Schooling? Evidence from Kenya. *Journal of African Economies*, 21(5): 691–722.

Ministry of Basic and Secondary Education, Gambia. 2004. *Sources of Funds for the Girls Scholarship Program*. Tech. rept.

Muralidharan, K., and N. Prakash. 2013. “Cycling to School: Increasing Secondary School Enrollment for Girls in India.” *Working Paper 19305*.

Muralidharan, K., and K. Sheth. 2015. “Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers.” *Journal of Human Resources*.

Murnane, R.J., and A. J. Ganimian. 2014. “Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations.” Working Paper 20284. National Bureau of Economic Research.

Petrosino, A., C. Morgan, T. A. Fronius, E. E. Tanner-Smith, and R. F. Boruch. 2012. “Interventions in Developing Nations for Improving Primary and Secondary.” *Campbell Systematic Reviews*, 19.

Pritchett, L. 2013. “The Rebirth of Education: Schooling Ain’t Learning.” Washington, DC: Center for Global Development.

Pugatch, T., and E. Schroeder. 2014a. “Incentives for Teacher Relocation: Evidence from the Gambian Hardship Allowance.” *Economics of Education Review* 41: 120–36.

———. 2014b. “Teacher Pay and Student Performance: Evidence from the Gambian Hardship Allowance.” IZA Discussion Paper 8621. Institute for the Study of Labor (IZA).

Sandefur, J., and A. Glassman. 2015. “The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics.” *The Journal of Development Studies* 51 (2): 116–32.

Webb, M. D. 2013. *Reworking Wild Bootstrap Based Inference for Clustered Errors*. Tech. rept. 1315. Queen’s Economics Department Working Paper.

World Bank. 2015. *World Development Indicators*. Tech. rept. World Bank.

World Bank. 2016. *EdStats*. Tech. rept.