

Online Appendix
Can Education be Standardized? Evidence from Kenya

Guthrie Gray-Lobe* Anthony Keats† Michael Kremer‡ Isaac Mbiti§
Owen Ozier ¶

Working paper
This version: June 6, 2022

*University of Chicago, email: graylobe@g.harvard.edu

†Wesleyan University, email: akeats@wesleyan.edu

‡University of Chicago, email: kremermr@uchicago.edu

§University of Virginia, email: imbiti@virginia.edu

¶Williams College and World Bank, email: owen.ozier@williams.edu

Contents

1	Construction of endogenous attendance variables	2
2	Complier baseline characteristics	4
3	Assessing the reliability of pupil identity information	4
4	Robustness of test score effect estimates	6
5	Early childhood literacy and numeracy items	8
6	Complier means by grade	9
7	Effects on the distribution of test scores	9
7.1	Kernel density plots of recipient and non-recipient test scores	9
7.2	Complier test score CDFs	10
7.3	Quantile regression results	10
8	Test score effect heterogeneity across pupil characteristics	11
9	Test score effect heterogeneity across classroom characteristics	12
9.1	Teacher characteristics	13
9.2	Class size	15
10	Heterogeneity in test score effects across Bridge locations	16
10.1	Test score effect heterogeneity across academy characteristics	16
10.2	Direct test for heterogeneity	18
11	Spoon task robustness	18
12	Effect of the scholarship on occupational aspirations	18
13	Effects on social attitudes	20
14	Effects on school characteristics and mediating inputs	21
15	Testing for an income effect of the scholarship	21
15.1	The effect of Bridge on non-recipient siblings	22
15.2	Robustness of results separately identifying the income and Bridge enrollment effects	23
16	Separate effect of Bridge enrollment and fee payment on timely grade progression	24
17	Additional results on the identification of sub-LATEs	24
17.1	Plug-in estimates	25
18	Non-experimental results on external validity	26
18.1	Data	27
18.2	Identification strategy: duration exposure	28

19 Bridge academy staff characteristics	29
20 Characteristics of Bridge pupils and classrooms	29

1 Construction of endogenous attendance variables

This subsection describes the procedure used to determine the type of school attended in 2017 by each pupil with endline data. Additional details can be found in [Gray-Lobe et al. \(2020\)](#).

Pupils are classified as having attended a Bridge school if their parents indicated that they attended a Bridge school (*pa_schltype_brg_studykid*). All pupils are classified as having attended public schools in 2017 if their parent responded that they attended a “government” school in 2017 (*pa_schltype_studykid*). Finally, a pupil is classified as attending a private school if their parent indicated that they attended a private school (*pa_schltype_studykid*), but did not indicate that they attended a Bridge school (*pa_schltype_brg_studykid*). 12,964 observations are classified using this approach.

If the caregiver indicated (*pa_att_school_2017*) that the pupil was not currently enrolled in school and the pupil survey (*pu_att_school_2017*) does not contradict this information (either the pupil doesn’t know or the item is missing), the pupil is classified as being unenrolled. Any pupil who reports that they were not enrolled in 2017 and have not been classified in a previous step are classified as unenrolled. 135 observations are classified using this approach.

Next, for any pupil who reports that the name of their school¹ contains the word “bridge” (*pu_school_contains_bridge* = 1) and who had not been classified by any previous step, we classified the pupil as having been enrolled in Bridge. 67 observations are classified using this approach.

Any pupil who was found to be attending a Bridge academy in the 2017 phone call survey (*pc17_schl_bridgelyn*=1) and who is not classified by the procedures above is classified as having attended a Bridge academy in 2017. 118 observations are classified using this approach.

Pupils are recorded as unenrolled if their parents indicated that they were not in school at the time of the survey (*pa_inschool_studykid*=0), they say they were not enrolled in their school for the full year of 2017 (*pu_school2017_allyear*=0) and the survey is missing data for the pupil’s response to the question of whether they attended school in 2017 (*pu_att_school_2017*). 12 observations are classified using this approach.

Twenty-eight pupils are classified on the basis of the 2016 phone call survey. Pupils who re-

¹This item asks the pupil what the name of the school is that they attended “in the last term of 2017”. For pupils who were unsure what time period was being referred to, the field officer would prompt that this was around the time of the recent August elections.

sponded to the 2016 phone call survey and who reported in the endline survey that they had not switched schools since 2016 were classified according to their response in the 2016 survey. Specifically, if a parent indicates that the student did not switch schools between 2016 and 2017 ($pa_m6_5=0$) we classify pupils as attending Bridge if their parents indicated they were attending Bridge in 2016 ($pc16_school=1$). Pupils are classified as having attended government school if the parent indicates that the student did not switch schools between 2016 and 2017 ($pa_m6_5=0$) and that they were attending a government school $pc16_govt_nongovt=1$ in 2016. Pupils are classified as having attended a private school if the parent indicates that the student did not switch schools between 2016 and 2017 ($pa_m6_5=0$) and their parents indicated that they were attending a non-Bridge private school in 2016 $pc16_govt_nongovt=2$ and that school was not a bridge school ($pc16_school=0$).

Two pupils are classified based on responses to a question to parents about the type of school into which the pupil switched between 2016 (year 1) and 2017 (year 2). Parents could report that their child switched into a public, private or Bridge school (pa_m6_7).² Parents also indicated the type of their school where their child had been attending prior to the switch (pa_m6_6). Parents who reported that their child had switched schools since 2016 into the same type of school they were in previously ($pa_m6_6 = pa_m6_7$) were classified as having been enrolled in 2017 in that type of school indicated in both fields.

Sixteen pupils are classified based on the name of the school that the pupil reports in the endline survey ($pu_school_likely = 1$). Schools are classified as being government schools if the name of the school includes the phrase “primary” and includes a reference to the physical location of the school. Schools are classified as “private” otherwise.³

Eight pupils are classified as being unenrolled if their parent responded that they were not in school at the time of the endline survey ($pa_in_school_studykid = 0$) as long as they were interviewed prior to January 2, 2018 ($do_fc(pa_submission_date) < td(02jan2018)$).

²The field officer noted that if they switched more than once, they should note the first school type in 2016, and the last school type in 2017

³This decision rule was adopted by observing that the word “primary” is highly associated with government school attendance. 91 percent of government school attendees reported the word “primary” in their schools’ names, compared to whereas 9 percent of non-government school attendees.

2 Complier baseline characteristics

Table 1 reports the complier means in the non-recipient samples and compares this to the full sample. The non-recipient complier mean is estimated using the approach of [Abadie \(2002\)](#) and is the 2SLS estimate of $1 - D_i$ on the transformed outcome $V_i^{1-D} = X_i(1 - D_i)$. The recipient mean is calculated as the 2SLS estimate of D_i on the transformed outcome $V_i^D = X_i D_i$.

Compliers are slightly less likely to be female than the full applicant sample. They tend to be more likely to have a dirt floor in their home and mud walls.

3 Assessing the reliability of pupil identity information

This section discusses the reliability of data identifying pupils in the application files that were received through staff at Bridge academies and how reliability of these data may influence the interpretation of the results. As noted, identities of those pupils who applied through staff at Bridge academies were not verified by the research team prior to the endline activity. This raises a concern that attempts to transfer the scholarships to non-recipients may compromise the results. We discuss first the problem itself as well as anecdotal and empirical evidence against the hypothesis that the estimated effects are influenced by scholarship transfers and that the extent of transfers must have been limited.

The rules of the scholarship program explicitly stated that the scholarships could not be transferred. These rules were communicated those staff at Bridge academies who would have been responsible for contacting prospective applicants. Transfer of a scholarship would require the fraudulent recipient to impersonate the intended scholarship recipient during interactions with both survey staff and employees of Bridge.

If scholarships were transferred to either non-recipient applicants or non-applicants, this would potentially bias the estimated treatment effects. Assuming that prospective scholarship buyers on the secondary market tend to be people who are more likely to attend Bridge anyway and that scholarship sellers tend to be those who are less likely to attend Bridge regardless of scholarship status, the secondary market would tend to increase the *observed* first stage effect of the scholarship on attendance compared to the first stage effect that one would observe in the absence of identity

fraud. This would attenuate the estimated 2SLS effect toward zero. Moreover, the ITT effect of the scholarship would be biased due to selection because buyers may be higher or lower performing on average.

We note several empirical results that support the view that transfer of scholarships was not widespread.

- Fraud would require coordination of several people, including children, and it is unlikely that this would escape detection. In practice, the field team did not report a single case of suspected identity fraud.
- We find positive effects in the P_{brig} and PP_{brig} samples. The identities of these applicants were verified by the fact that all pupils were previously enrolled in Bridge schools. The research team confirmed this fact by matching applicants to administrative enrollment records provided by Bridge.
- The estimated effects on the sample of pupils who applied through IPA are similar in magnitude. The identities of these applicants were verified through interactions with IPA staff prior to scholarship assignment. All pupils who applied through IPA were required to sit for baseline exams.
- The overall follow-up rate is above 85 percent in the main samples, and the attrition differential is only one percent overall. If scholarship buyers must abandon their original applications, as would be necessary in most cases to escape detection by the field team, then this differential suggests that the scope for widespread *buying* of scholarship on the secondary market is unlikely (although it is still possible that there was widespread *selling* of scholarships on the secondary market to non-applicants).
- Non-experimental duration exposure estimates reported in Section 18.2 also show large positive effects (although the units of the scores cannot be compared directly to those used in the main analysis).
- An experimental strategy using scholarship assignment among siblings finds similar point estimates (although the estimated effects are not statistically significant, due to low precision of the test).

In considering the scope for illicit transfer of scholarships we have mostly considered the possibility that fraudulent recipients impersonate the original recipient in order to assume their identity. However, we do note that in some cases, names and characteristics of pupils could be updated by field officers if an error was found. In some cases, mistakes in the application data were detected. A common mistake was for the person completing the application to use their own name instead of the pupil’s name. Table 2 reports the ITT effect of measures of the distance between baseline and endline characteristics. These include the Levenshtein string distance between the name of the pupil, the age distance, and an indicator for whether the genders match. In the case of age distance, the application includes the year of birth. Endline data include the pupil’s age in years. Their baseline age is computed as (2017 - baseline year of birth). Overall, we find the baseline and endline characteristics are highly correlated. The average pupil name is less than one character off from baseline to endline. Pupils on average are 1.01 years from their baseline projected age. 97 percent of pupils in the main samples have the same gender in the baseline and endline data. Overall, we find that the scholarship has little or no effect on these measures, especially in the main samples. We do find that ages are more similar in the PP_{main} . While this may indicate some effect of the scholarship on reported ages, it is inconsistent with the view that scholarships were being transferred on a secondary market.

4 Robustness of test score effect estimates

This section presents robustness checks on the main results to alternative data processing choices. We examine robustness to applying different inclusion criteria to form the main analysis sample, to different choices of standard deviation units, and to alternative methods for equating the academic test score results.

The results are robust to restricting the analysis to the sub-sample of pupils who submitted applications through IPA (see Table 3). We view this sample as a having high quality application data because the identities of the applicants were verified by staff of a third-party research organization. It is reassuring that the results in this sample are broadly similar to those in using the full sample.

The results are robust to expanding the sample to include all applications from Meru and

Nakuru counties (see Table 4). As noted in the text, we are concerned about the quality of these data, but it is reassuring to note that the results are not meaningfully changed by the decision to exclude them in the main analysis.

The results are robust to the equating method used to place performance on grade-specific tests on a common scale. As discussed in the paper, academic subject tests at endline were designed to reflect content that children were learning in school in their current grade. In the main results, we placed grade-specific test scores on a common scale using Item Response Theory (IRT) equating methods. Specifically, we used the Stocking-Lord method. Table 5 reports results using alternative equating methods. For discussion of individual equating methods, see [Kolen and Brennan \(2014\)](#).

The numerical values of the estimated effects depend on the population standard deviation in which the test scores are expressed. Table 6 illustrates the sensitivity of the numerical values to decisions related to the population from which the standard deviation units are drawn. Columns 1 and 7 of Table 6 show the effects of enrolling at Bridge in the units that were specified in [Gray-Lobe et al. \(2020\)](#): standard deviations of the Standard (grade) four test score distribution. Columns 2 and 8 express results in standard deviation units of the pupil’s projected grade. For a pupil who was entering grade g in 2016, this is the standard deviation of all pupils entering grade g . Because baseline entering grade appears to have contained some measurement error so that the cohort of children who, according to application data, were entering grade g includes some children who were actually entering different grades, the standard deviation units reported in Columns 2 and 8, restrict to those pupils who took the correct form. Columns 3 and 9 include all pupils who were entering grade g . Columns 4 and 10 report results from a coarser approach to standardizing test scores within baseline cohorts where the standard deviations are either for the full pre-primary or primary sample, depending on the grade group where the applicant was entering in 2016. Columns 5 and 11 report results where the test scores are standardized across the entire distribution of applicants. Finally, Columns 6 and 12 report results where we do not equate the test scores and leave the scores in standard deviation units of the grade where the pupil was found at endline.

Table 7 shows the results where we replace Bernoulli Bridge enrollment indicator with a count of the number of years that the pupil was enrolled at Bridge. The 2SLS estimates are the average causal response of a year of enrollment at Bridge ([Angrist and Imbens, 1995](#)).

5 Early childhood literacy and numeracy items

For younger pupils it is straightforward to link individual items to fundamental literacy and numeracy skills. Therefore, for the sample of pupils who were planning to enroll in pre-primary grades in year 1, we examine the effects of the scholarship and Bridge attendance on individual items. For this analysis, we also present results for individual grade levels to facilitate comparison to other studies on early childhood development.

Literacy We measure specific literacy skills using three items. “Letters per minute” indicates the number of correct letters identified in one minute. For this task, pupils were shown 60 letters that included both upper and lower case letters. Some letters were repeated. Only 1.3 percent of pupils were able to identify 60 letters in 60 seconds. Ability to “read a simple sentence” was measured by asking the pupil to read the sentences “Ken has a big dog” and “Kaka alikunywa sharubati”. We also measure reading ability with a question that asked pupils to read the word “Thursday.”

Overall, attending Bridge appears to increase all three skills (Table 8). While scholarship non-recipient pre-primary pupils were able to identify 15.8 letters per minute, attending Bridge increases the number of letters identified per minute by 11.87. Attending Bridge increases the probability that a pupil can read the word “Thursday” by 14 percentage points, an effect size that itself is more than double the non-recipient mean. This effect is, however, concentrated among the older pupils who were projected to be in Standard 1 at endline. Attending Bridge also increased the likelihood that a pupil could read a simple sentence by 31 percentage points, compared to 22 percent of non-recipients. For Standard 1 pupils, attending Bridge increased the probability that a pupil could read the simple sentence by 54 percentage points.⁴

Numeracy We measure basic numeracy using four items. First, pupils were shown eight stars and asked “how many stars are here?” After responding to that question, pupils were asked “How many stars are left if you take away 5 stars?” These questions were only asked of pupils who were enrolled in ECD grades at endline. Two questions asked of both ECD and Standard 1 students are

⁴Note that this estimate does not mean that the predicted treatment effect is outside the unit interval, violating the assumptions of the linear probability model. Using the technique of [Abadie \(2002\)](#) we estimate the predicted outcome for compliers who attend Bridge as 0.82, inside the unit interval.

also included. First, we report results on a question that asked pupils to add “4+2.” Second, we report results on a question that asked pupils to complete a sequence of numbers “9,8,7,-”.

For each of the numeracy items, attending Bridge increases the probability of a correct response (Table 8). Attending Bridge increases the likelihood that a pupil can count eight stars by 12 percentage points, can subtract five from eight stars by 13 percentage points, and can add “4+2” by 20 percentage points. Attending Bridge increases the probability that a pupil can complete the sequence “9,8,7,-” by 12 percentage points, from a base of 19 percent.

Table 8 presents results that were not pre-specified in the PASPA. The research team selected items that would facilitate comparison to other early grade literacy and numeracy assessments. We find that the effects on these particular test items are not outliers when compared to other items on the assessments.

6 Complier means by grade

Figure 1 graphs the estimated means of potential outcomes in Bridge, $Y(\text{Bridge})$, and not in Bridge, $Y(\text{Not Bridge})$, for compliers separately by projected year 2 (2017) grade using the approach of [Abadie \(2002\)](#). $Y(\text{Bridge})$ is the 2SLS estimate of attending Bridge $D_i \in \{0, 1\}$ on the transformed dependent variable $W_i^D = Y_i D_i$. $Y(\text{Not Bridge})$ is the 2SLS estimate of $1 - D_i$ on the transformed outcome $W_i^{1-D} = Y_i(1 - D_i)$.

7 Effects on the distribution of test scores

This section describes additional results related to the effect of the scholarship and of Bridge enrollment on the distribution of test scores.

7.1 Kernel density plots of recipient and non-recipient test scores

Figures 2 and 3 show kernel density plots of the test scores distribution for scholarship recipients and non-recipients in the primary (P_{main}) and pre-primary samples (PP_{main}), respectively. The standard deviations and an F-test of equality of the standard deviations is reported below the figures.

7.2 Complier test score CDFs

Figures 4 and 5 shows the estimated distribution of test scores for scholarship recipients and non-recipients in the primary and pre-primary samples respectively using an approach from [Abadie \(2002\)](#). Each point in the estimated CDF reflects a separate 2SLS estimate. For recipient CDF gives the 2SLS estimated effect of attending Bridge $D_i \in \{0, 1\}$ on the transformed dependent variable $W_i^D = \mathbb{1}\{Y_i > x\}D_i$ where x represents a point in the support of the test score distribution. The non-recipient CDF gives the 2SLS estimated effect of not attending Bridge $1 - D_i$ on the transformed outcome $W_i^{1-D} = \mathbb{1}\{Y_i > x\}(1 - D_i)$.

7.3 Quantile regression results

Figure 6 shows the results of IV quantile regression using the estimator from [Abadie et al. \(2002\)](#). Figure 7 shows analogous results using the estimator from [Chernozhukov and Hansen \(2013\)](#). The results indicate positive effects across at all quantiles. The results using [Abadie et al. \(2002\)](#) estimate a *local* quantile treatment effect under the standard assumption of monotonicity whereas the results from [Chernozhukov and Hansen \(2013\)](#) estimate a global quantile treatment effect under an assumption of “rank similarity”.⁵⁶ Table 9 shows the results using the estimator proposed by [Abadie et al. \(2002\)](#).

The results using both methods are broadly similar. Both approaches show statistically significant positive effects at all quantiles in the distribution. The estimates using [Chernozhukov and Hansen \(2013\)](#) indicate a clear pattern of larger effects for lower quantiles. While there may be a similar pattern using the estimates from the [Abadie et al. \(2002\)](#) approach, the pattern is less clear.

⁵See [Wüthrich \(2020\)](#) for a comparison of the two estimators.

⁶The PAPSA discussed the use of [Chernozhukov and Hansen \(2013\)](#). However, because the assumptions invoked for [Abadie et al. \(2002\)](#) are more similar to those under conventional IV estimation in the analysis of encouragement experiments, we have focused on those results.

8 Test score effect heterogeneity across pupil characteristics

This section tests for heterogeneous effects of attending Bridge by estimating the following structural model

$$y_i = \alpha_1 D_i + \alpha_2 D_i \times C_i + X_i \Gamma + \epsilon_i , \quad (1)$$

where y_i is the aggregate subject knowledge index, $D_i \in \{0, 1\}$ indicates attendance at a Bridge school, C_i is an individual level covariate, X_i is a vector of control variables, including strata dummies, and ϵ_i is an individual level disturbance term. To estimate this model, we interact treatment with the covariate to form a second instrument so that the model is just identified.

We test for heterogeneous effects using several covariates including gender and measures of baseline achievement. We provide additional information measures of baseline achievement used to form the interaction terms below. The results are presented in Table 10.

Gender. The point estimate on the interaction term with an indicator for female is negative, indicating that girls may benefit less than boys, but these results are not statistically significant. Table 11 further shows separate effects by gender on other outcomes.

Math and English percentile ranks. The application data include self-reports of the pupil’s end-of-term tests in English and math for the school year prior to the two-year scholarship lottery. These are not standardized tests, but pupils within the same geographic area in the same grade may have taken similar tests. Within each grade-academy cell of the P_{main} and PP_{main} sample, we form the percentile rank of pupils using their reported math and English scores to form a math percentile and an English percentile score.⁷ The self-reported test score outcomes are only available for pupils who report them, so there is some loss of sample size for this test.

For English percentile rank, the interaction term is negative and statistically significant, indicating that higher achieving pupils benefit less from attending Bridge. The point estimate of for the interaction term with the math percentile is also negative, but not statistically significant.

⁷We find that the transformed math and English scores are more predictive of their endline counterparts than the reported percentage correct scores.

Baseline cognitive and non-cognitive skills. Self-regulation, the Malawi Development Assessment Tool (MDAT), and Raven’s score baseline data were only collected for a sub-set of pupils whose families applied through IPA. Individual assessments were administered only to certain age groups. Only primary school-aged pupils received the Raven’s score baseline test. The self-regulation test was administered to all pupils. The point estimate for the interaction with the self-regulation scores is negative, but these effects are not statistically significant. The point estimates for the interactions with the MDAT scores is positive, but these results are not statistically significant.

Expected achievement index The statistical power of the tests above is limited by the fact that the availability of individual measures of baseline pupil achievement is not consistent across age groups. To obtain an estimate of heterogeneous effects of Bridge across pupils with with varying baseline achievement, we use available baseline characteristics to form a prediction of endline test scores. This procedure generates predicted outcomes by estimating a model of the outcome as a function of baseline covariates using the full sample of experimental controls and then using the estimated model, they predict outcomes for the full sample.⁸ Baseline controls used to construct the predicted endline test score are the self-reported term 3 end-of-term English and math scores, scores on individual cognitive and non-cognitive assessments, pupil height, pupil gender, a variable indicating whether the primary caregiver is literate, and indicators for whether the primary caregiver is engaged in agricultural labor, casual labor, is a small business owner, or unemployed. Where data are missing for individual pupils, we impute the missing variable to the mean and include a dummy variable indicating that the value was missing, and baseline projected 2016 grade.

9 Test score effect heterogeneity across classroom characteristics

This section tests for heterogeneity in effects across characteristics of Bridge classrooms. Section 9.1 tests for heterogeneity in the effect of attending Bridge when Bridge teachers have different levels of experience, education, and subject knowledge. Section 9.2 tests for heterogeneity in the

⁸Abadie et al. (2018) show that this strategy can lead to bias due to mean-reversion in the treated group. However, they present evidence that leaving individual i out of the model used to form i ’s prediction can remove this bias. We employ this strategy to form a leave-one-out (LOO) predicted subject knowledge index.

effects of Bridge across classrooms with different class sizes on pupil reported measures of classroom disruptions, adaptive instruction, access to textbooks, and test score outcomes.

9.1 Teacher characteristics

This section examines whether characteristics of Bridge teachers correlate with the size of the effect of attending Bridge on subject knowledge test scores. We estimate the following structural model for a pupil i 's test score in subject s :

$$y_{i,s} = \alpha_1 D_i + \alpha_2 D_i \times C_{s,g(i)} + \alpha_3 C_{s,g(i)} + \epsilon_{i,s} , \quad (2)$$

where $C_{s,g(i)}$ represents a characteristic of a bridge teacher in pupil i 's 2017 grade level $g(i)$. Subject scores in English, Kiswahili, and math are stacked and errors are clustered at the pupil level. We estimate Equation 2 using 2SLS with two instruments: scholarship assignment Z_i and an interaction of Z_i with the characteristic of the Bridge teacher for the grade in which the pupil was projected to enroll in 2017 $C_{s,g(i)}$.

More precise estimates of the effect of being enrolling in Bridge classrooms with different characteristics can be obtained using a pupil fixed effect model similar to [Bold et al. \(2017\)](#) under stronger assumptions. Specifically, we estimate the following pupil fixed effect model for pupils enrolled in Bridge in 2017:

$$y_{i,s} = \alpha + \beta C_{s,g(i)} + \lambda_i + \epsilon_{i,s} , \quad (3)$$

where λ_i is a pupil fixed effect, and $\epsilon_{i,s}$ is a subject-specific error term. Under an assumption that the subject specific error term is uncorrelated with $C_{s,g(i)}$, $E[\epsilon_{is}|C_{s,g(i)}] = 0$, $\hat{\beta}$ can be interpreted as an estimate of the effect on a pupil outcome of receiving a classroom with 1 unit more of $C_{s,g(i)}$. For these tests we use the unequated IRT test score because the identification is from variation between subjects within the same pupil. As a placebo test, we also estimate the same model for the sample of pupils who were not enrolled in Bridge in 2017. The pupil fixed effect estimates implicitly restricts to upper primary students because lower primary and pre-primary classrooms do not have variation in teacher characteristics because a single teacher teaches all subjects.

Characteristics of Bridge teachers are measured using administrative data provided by Bridge.

Teacher data are linked to Bridge classrooms using a file that contains only primary school grades, and therefore pupils who are projected to be in pre-primary grades in 2017 are not included in any tests that use these interactions. This restriction eliminates all but the oldest cohort in the pre-primary sample.

The characteristics of teachers used to form the interaction terms are described below.

Experience, Certification and KCSE score. Teacher experience is measured using Bridge data on hiring and separations. Bridge administrative data include the year of hire which is used to identify years of experience teaching *at Bridge* for the teacher instructing a given subject for a given grade.⁹ Certification and KCSE scores are measured using recruitment and training data files maintained by Bridge.

Discussion. At the primary school level, less experienced teachers are associated with smaller effects, but the interaction term is not statistically significant. First year teachers are associated with 0.10 standard deviation smaller effects in primary school (Table 12). At the pre-primary level, the sign of the interaction term is reversed and marginally statistically significant. The pupil fixed effect estimates are more precise and small in magnitude, suggesting that teacher experience is not predictive of classroom effects.

At the primary school level, certified teachers are associated with 0.10 standard deviation smaller effects in primary school. The Pupil FE estimates also show that certified teachers have smaller effects and this estimate is marginally statistically significant. We note that this result means that certified teachers may produce slightly smaller effects than non-certified teachers. This could reflect differences in experience with the Bridge model because more recent hires have been more likely to be certified (Figure 10), challenges due to retraining for a unique classroom environment, or simply sampling error.¹⁰ It is also possible that the pool of TSC certified teachers who are willing to work at Bridge may be negatively selected relative to the wider pool of secondary school

⁹Data on prior experience teaching is unavailable except for teachers hired during 2016, so it is not possible to identify teachers in their first year of teaching in any classroom. However, the data indicate that 99 percent of years hired in 2016 had no prior experience teaching.

¹⁰Bridge's hiring practices appear to have changed markedly around the time of the study, and these changes may influence the interpretation of these results. Figure 10 shows that, in recent years, Bridge has begun hiring a greater share of teachers with certification. Therefore, more recent hires are also more likely to have certification. The strong negative relationship between teacher certification and the number of years of experience teaching at Bridge is also illustrated in the correlation matrix in Table 13.

educated people because more educated teachers tend to have better outside options. Importantly, the magnitude of the interaction effect is small and accounts for a small share of the total effect.

The absence of a strong correlation between observable Bridge teacher characteristics and the effect of attending Bridge is consistent with the view that teachers with varying levels of education, experience, and subject knowledge can deliver similar effects using Bridge’s materials. Under this interpretation, Bridge’s standardization of lesson planning may substitute for gaps in skills among less educated teachers. The results from the interaction of first year teaching may be the especially surprising given a large literature documenting lower teacher value-added of first year teachers (Kane et al., 2008; Chetty et al., 2014; Staiger and Rockoff, 2010). However, we note that this interpretation of these results as evidence that Bridge’s teacher guides substitute for teacher skills assumes that these characteristics would be predictive of classroom value-added in other settings. Importantly, our data are silent on these characteristics in local public schools and non-Bridge private schools.

9.2 Class size

This section tests for heterogeneity in the effect of attending Bridge on test scores across smaller and larger classrooms.¹¹ Before examining the relationship between class size and test score outcomes, we examine the relationship between class size and intermediate outcomes related to classroom inputs such as the number of textbooks a pupil reports having, the classroom crowding and classroom engagement indices, and the adaptivity index. Results are presented in Table 14.

Discussion. The estimated effect of attending Bridge is similar across smaller and larger classrooms on measures of classroom inputs and on test scores, and the results are consistent with the view that even larger Bridge classrooms also generate positive effects on test scores for both primary and pre-primary students. However, we cannot rule out the possibility that larger classrooms would have substantially smaller effects when we extrapolate to class sizes comparable to public schools.

¹¹The number of pupils in a classroom is calculated as the number of pupils in a grade during the first term of 2017 according to internal test score data supplied by Bridge. Because all Bridge academies had a single stream (a group of students who receive instruction together), the number of pupils per grade is coincides with the number of pupils in a classroom and the pupil-teacher ratio. According to this measure, over half of pupils attending Bridge are in classrooms with fewer than 20 pupils.

At the primary school level, point estimates for the interaction term indicate larger class sizes attenuate the effects on access to textbooks, classroom crowding, teacher-classroom engagement, parental engagement, instructional adaptation, and test scores, but we cannot reject the null hypothesis for any of these outcomes (Table 14). However, we cannot rule out the possibility that Bridge’s effect would be substantially smaller if its classrooms contained as many children as public schools. Although the difference is not statistically significant, these results indicate that Bridge’s effect on a classroom with 43 students (the average public school class size) would be 0.270 standard deviations smaller than that in a classroom with 13 students, (the average Bridge school class size). This is about one third the effect in the average Bridge classroom.

At the pre-primary level, the sign of the effects on the measures of inputs are not consistent with the view that increasing class size reduces per pupil resources and the estimated effect of larger classrooms is positive, and we cannot reject the null hypothesis for any of these outcomes.

Although we cannot rule out the possibility that the effect of Bridge would be smaller in larger classrooms, these results are inconsistent with the view that the effect of Bridge on test scores is explained primarily by the large difference in class size. However, we note that the absence of a correlation between Bridge class size potentially reflect correlation of class size with unobserved characteristics, including characteristics of the counterfactual learning environment. If, for example, large class sizes in Bridge schools are mirrored by large class sizes in local public schools, then the interaction effect may understate the effect of a *ceteris paribus* change in class size.

10 Heterogeneity in test score effects across Bridge locations

This section tests for differences in site effects across Bridge academies with different observable characteristics and reports the results from a over-identification test of homogeneous site effects.

10.1 Test score effect heterogeneity across academy characteristics

To assess whether the estimated effect of attending Bridge varies systematically across sites according to observable characteristics, we estimate the following model

$$y_i = \alpha_1 D_i + \alpha_2 D_i \times C_{ia} + X_i \Gamma + \epsilon_i$$

where y_i is the aggregate subject knowledge index, $D_i \in \{0, 1\}$ indicates attendance at a Bridge school, C_{ia} is an academy level covariate, X_i is a vector of control variables, including strata dummies, and ϵ_i is an individual level disturbance term. To estimate this model, we interact treatment with the covariate to form a second instrument so that the specification is just-identified.

Table 15 shows the results of tests of heterogeneous effects across a variety of academy level covariates. We discuss each of the characteristics interacted with Bridge attendance below.

Local public school quality. Table 15 includes several measures of the quality of local public schools. First, using data on KCPE scores from 2013, we form the county mean KCPE score and link this characteristic to the academy. Second, we identify the nearest government school with 2013 KCPE using a census of schools that contains GPS locations. After matching schools geographically, the median KCPE score of pupils in the nearest government school is formed. Although this more local measure of school quality may be superior to the county average, using only a single school may induce measurement error if this approach cannot precisely identify the government school that a pupil in the scholarship program would have gone to in the absence of the scholarship, especially in urban areas where there may be several nearby government schools. With this in mind, we form a third measure of local school quality using the mean KCPE score of all government school pupils within two kilometers of the Bridge school.

Academy age. Bridge schools may require some time to develop systems locally to be effective. We test whether Bridge schools that have been operating for longer are more effective by interacting Bridge attendance with the number of years of operation of the Bridge school. Data on the first year of academy operation were provided by Bridge.

Rural/Urban. Bridge currently operates across a broad range of geographic areas, including both rural and urban areas. Bridge provided classifications of all academies in as either rural, urban, or peri-urban. This classification is maintained in part for the determination of cost of living adjustments for salaries.

Community socioeconomic status. Data on household income of applicants to Bridge come from the baseline application. These are used to calculate the mean household income of families

that applied to each academy.

10.2 Direct test for heterogeneity

This section tests directly for heterogeneity in effects across locations. We report the results of an overidentification test that instruments Bridge attendance with interactions of treatment with location dummies. We report the Sargan statistic and p-values in Table 16.

For these tests, we restrict the sample to academies with more than 10 observations. Simulation suggests that when location cells are small, the distribution of the Sargan statistic tends to over-reject the null hypothesis.

All of the tests fail to reject the null hypothesis of no heterogeneity in effects.

11 Spoon task robustness

Table 17 shows the effect of the scholarship and Bridge enrollment on outcomes related to the spoon use task. The spoon use task asked children to enumerate as many uses of a spoon as they could. During the interview field officers recorded their subjective judgment about whether each use was “unusual”. The task was then scored by summing the number of responses (fluency) and the number of unusual responses (originality). To evaluate the sensitivity of the scoring to subjective field officer judgments, a later activity had external judges on mTurk classify responses as either unoriginal, somewhat original, or very original. Each item was rated by over 240 individual judges. At the item level, we calculate the fraction of judges who assessed the response to belong to each category. We then calculate an analogous originality score by summing the shares saying that the response was either somewhat or very original. The point estimates are similar regardless of whether we use the field officer or external judge classifications, although estimates are more precise when using the external judges.

12 Effect of the scholarship on occupational aspirations

This section estimates effects of the scholarship and of attending Bridge on the occupational aspirations of pupils. Pupils were asked what they would like to be when they grow up. Field officers

classified occupations into one of 50 job categories. 5.0 percent of respondents did not know or refused to answer this item. 3.6 percent reported an occupation that was not on the list.¹²

Occupational indices. We collapse the high dimensional occupational aspiration outcome into a single variable by constructing indices that relate occupational aspirations to pupil characteristics. These indices can be interpreted as leave-‘i’-out predictions of a characteristic θ_i , where θ_i is either the pupil’s endline aggregate subject knowledge index or an indicator for the pupil being female. In each case, we estimate the OLS relationship between θ_i a vector of dummy variables O_i representing each occupation.

$$\theta_i = \alpha + O_i\Omega + \epsilon_i . \tag{4}$$

We then use the estimated model to form the leave-i-out prediction $\hat{\theta}_i^{LOO}$.

Occupation indicators. We also examine the effect of the scholarship on aspirations to specific occupations or groups of occupations. First, we group all occupations according to whether they are “professional” occupations that would typically require advanced training.¹³ Additionally, we examine effects on the largest occupation categories. These include doctor, teacher, engineer, pilot and driver.

Overall, attending Bridge does not appear to have any significant effect on occupational aspirations (Table 18). The one exception may be that older pupils in the P_{govt} sample seem to aspire to occupations associated with higher achieving pupils, but this effect is small and only marginally significant.

Joint test of any effect on occupational aspirations. We test the null hypothesis that scholarship receipt had no effect on aspirations for any particular occupation. To do so we regress the scholarship offer on indicators for all occupational aspirations and randomization strata dummy controls. The results of this test are reported at the bottom of the table.

¹²“Other” responses have not yet been classified.

¹³We include as professional occupations accountant, civil servant, computer programmer, dentist, district collector, doctor, engineer, manager, journalist, lawyer, lecturer, nurse, politician, scientist, teacher, and veterinarian. We also include those who stated they aspired to be a university student because this likely indicates an interest in having a professional occupation. Some occupations, like politician or manager, may not require advanced degrees, but we feel that aspirations to these occupations will tend to coincide with aspirations toward advanced degrees.

Discussion. Among the older primary school cohorts, there is some evidence that attending Bridge induces aspirations toward the types of occupations preferred by higher performing pupils (*Occupation academic achievement*). However, for both pre-primary and primary school children, we fail to reject the null hypothesis of the joint test.

13 Effects on social attitudes

This section presents results on the effects of attending Bridge on pupils' social attitudes. The pupil survey contained a brief module on social attitudes, including inter-ethnic trust, support for democracy, and gender. Because effects on gender attitudes may differ across pupils of different genders, we report results pooling all genders and separately for boys and girls. We describe each of the outcomes below. Results are reported in Table 19.

Trust. Pupils were asked “in general, can you trust members of your tribe?” and “in general, can you trust members of other tribes?”. The item allowed for a “yes/no” response. For each question, 6 percent of pupils responded that they did not know or refused to answer. These responses are treated as missing.

Support for democracy. Pupils were asked if they agreed or disagreed with the statement that elections are the best way to choose leaders. Allowable responses were “strongly agree”, “agree”, “disagree”, and “strongly disagree”. We report results for an outcome that transforms responses on this scale into a binary outcome indicating whether the pupil agrees with the statement at all. We find similar results using ordered logit specification.

Gender. Pupils were asked if they agreed or disagreed with the statement “Boys are smarter than girls.” Allowable responses were “strongly agree”, “agree”, “disagree”, and “strongly disagree”. We report results for an outcome that transforms responses on this scale into a binary outcome indicating whether the pupil agrees with the statement at all. We find similar results using ordered logit specification.

14 Effects on school characteristics and mediating inputs

This section presents results on individual measures of school characteristics and mediating inputs that were not reported in the main document. The main document reported results on indices that were composed of these items, but the indices were constructed after observing effects on individual items. Some of the individual items were reported in the main analysis but are reported again to facilitate comparisons.

Table 20 reports characteristics of Bridge, public, and other private schools in the non-recipient sample. Table 21 shows the 2SLS effect of the scholarship on pupil reported characteristics of school facilities. Table 22 reports pupil reported effects on items related to instructional time including the length of the school day and Saturday school attendance, as well as teacher absences. Table 23 reports results on crowding and other classroom disruptions. Table 24 reports on the effects of attending Bridge on access to textbooks. Table 25 reports on the effects of attending Bridge on pedagogy, engagement, and feedback. Table 26 reports on test preparation, homework, and usual study hours. Table 27 reports on the effect of attending Bridge on measures of parental engagement. Table 28 reports on the effect of attending Bridge on corporal punishment. Table 29 reports on the effects of scholarship and Bridge attendance on test conditions.

15 Testing for an income effect of the scholarship

This section presents additional tests of a direct effect of the scholarship through its effect on household budgets. First, we report the results using random variation in scholarship receipt for non-recipient siblings. We show that the scholarship increased enrollment of non-recipient siblings. However, because the scholarship did not lower the price of enrollment at Bridge for these siblings, this source of randomly induced Bridge enrollment is arguably less compromised by income effects. Then we also report the results from additional specifications that estimate the separate effect of income and enrollment effects.

15.1 The effect of Bridge on non-recipient siblings

Many applicants live in households with other applicants. This section documents that the receipt of a scholarship had spillover effects on pupils within the same household who did not receive a scholarship.¹⁴

For each non-recipient with siblings, we calculate the average treatment status among siblings,

$$\zeta_i = \frac{\sum_{k:h(i)=h(k),i \neq j} Z_i}{N_{h(i)} - 1},$$

where $h(i)$ be a function that identifies the household associated with applicant i and $N_{h(i)}$ is the number of other applicants in applicant i 's household.

Defining ζ_i^{25} to be an indicator for whether more than 25 percent of pupil i 's siblings received the scholarship, a propensity score control can be obtained by simulating the randomization of scholarships S times and then estimating the frequency with which ζ_i exceeds 0.25.¹⁵

The effect on pupil i of offering a scholarship to siblings in the household $k \neq i$ is estimated by

$$D_i = \pi_0 + \pi_1 \zeta_i^{25} + \pi_2 P_i(\zeta_i^{25}) + \epsilon_i, \quad (5)$$

where $P_i(\zeta_i^{25})$ represents the propensity score and D_i is any endogenous attendance variable. Predicted attendance \hat{D}_i from estimating Equation 5 can then be used to estimate second stage of a 2SLS model:

$$Y_i = \beta_0 + \beta_1 \hat{D}_i + \beta_2 P_i(\zeta_i^{25}) + \eta_i. \quad (6)$$

Table 30 presents the baseline balance table for the sibling instrument design. The point estimates of the effect of attending Bridge using the sibling instrument (Table 31) are positive and similar in magnitude to the effects using individual scholarship assignment, but the results are not statistically significant. The sibling instrument increases the probability that a pupil in primary

¹⁴We use phone numbers and contact information to identify children linked by the same adult caregivers. Prior to endline activities, households were identified using caregiver contact information for logistical purposes, and the procedure was found to be accurate at predicting household membership. Seventy-one percent of applicants in the study are found to be in a household with another applicant using this method. Over 90 percent of these households have 5 or fewer applicants.

¹⁵One could create many instruments corresponding to different thresholds (50%,75%) with corresponding propensity scores and use these instruments in an overidentified 2SLS specification. Doing so has a negligible effect on precision.

school attends Bridge by 10 percentage points and increases the probability that a pupil in pre-primary school attends Bridge by 11 percentage points. Using this instrument, the 2SLS effect of attending Bridge on the primary school academic subject knowledge index is 0.60 and the effect on the pre-primary academic subject knowledge is 1.63 standard deviations.

Table 31 also reports the effects of attending Bridge using the sibling instrument on educational expenditures and missed class. Households spend more on total educational expenditures if at least 25 percent of siblings receive the scholarship. In the primary school sample, some of this increased expenditure appears to be accounted for by additional spending on the non-recipient pupil. Results on missed class are mixed. However, it is noteworthy that the estimated effects on missing class due to missed fees is positive for both samples.

15.2 Robustness of results separately identifying the income and Bridge enrollment effects

In this section we report additional specifications that separately identify the income effect and the effect of enrolling at Bridge. In the paper we separately identify the effects of paying fees and enrolling at Bridge using an overidentified model with instruments formed by interacting scholarship assignment with baseline covariates. In [Gray-Lobe et al. \(2020\)](#) we discussed the use of multiple sets of instruments to assess robustness and because the underlying assumption of homogeneity of the average causal response across pupils in different covariate cells could be more or less plausible, depending on which baseline covariates are used to form the instruments. Furthermore, [Gray-Lobe et al. \(2020\)](#) discussed estimation in different samples. In addition to the two sets of covariates discussed in the paper, we also report results using interactions with indicators for the applicants county of residence. We also discussed combining all covariates, including the 10 demographic and intention covariates, three academy-intention means, and the county dummies. Because this specification potentially is vulnerable to bias due to having many weak instruments, we also report results using LIML.

Table 32 reports results for the primary school sample of the separate effects of Bridge enrollment and fees paid. Panel A reports results for the P_{main} sample alone. Panel B reports results from pooling the P_{main} and P_{brig} samples. Panel C reports results from pooling the P_{main} , P_{brig} , and P_{priv} samples. Table 33 reports analogous results for the pre-primary sample. Panel A reports results

for the PP_{main} sample. Panel B reports results for the combined PP_{main} and PP_{brig} samples.

One concern with the results above is that if the first stage effect of Bridge enrollment is misidentified, this could produce a false appearance of heterogeneity in the effect of Bridge enrollment, and this false heterogeneity would be projected onto either the enrollment effect or the fee paying effect. The results estimate the effect of being enrolled in Bridge in 2017. As noted in the paper, Bridge enrollment in 2017 is highly collinear with enrollment in 2016. However, in the sample of current Bridge students, this is not the case. The effect on Bridge enrollment widens between 2016 and 2017 as some non-recipients leave Bridge. If the effect of Bridge scales linearly with the number of years exposed then using 2017 enrollment to estimate the first stage impact on enrollment could lead us to underestimate the effect of Bridge enrollment in the current Bridge sample and, because this sample experienced a larger reduction in fees as a result of the scholarship, the lower 2SLS effect on this sample lead us to overestimate the effect of paying fees.

One strategy to address this issue is to use the number of years enrolled in Bridge instead of the Bernoulli Bridge enrollment in 2017. Table 34 reports results for the primary school sample. Table 35 reports results for the pre-primary sample.

16 Separate effect of Bridge enrollment and fee payment on timely grade progression

Table 36 reports the results of estimating the separate effects of Bridge enrollment and the payment of fees – as in Section 15.2 above – on the timely grade progression outcome.

17 Additional results on the identification of sub-LATEs

This section provides additional results using the multiple endogenous variables framework to identify separate sub-LATEs associated with different counterfactuals. The instruments used to identify each specification are described in the main analysis.

Table 37 shows results that decompose the effect on pre-primary and primary grade students into an effect from attending public and private schools. Panel A reports results for the pre-primary sample. Panel B reports analogous results for the primary grade sample. Panels C and D add the

P_{priv} and brig samples to improve precision. Table 38 shows analogous results where the endogenous variables are constructed to represent the number of years that a pupil was enrolled in each school type.

17.1 Plug-in estimates

The estimate from 2SLS of the effect of attending Bridge relative to a mixed counterfactual ($\hat{\rho}$), is composed of a weighted average of the effects of attending Bridge relative to each of the counterfactuals where the weights are composed of the effects of the scholarship offer on attendance at each of the counterfactual environments.

$$\hat{\rho} = - \sum_{j \in \{G, P, U, B\}} \frac{\hat{\pi}_j}{\hat{\pi}_S} \rho_j ,$$

where $\hat{\pi}_S$ is the effect of the scholarship offer on attendance at Bridge, $\hat{\pi}_j$ indicates the first stage effect of the scholarship offer on attendance at each of the educational environments $j \in \{G, P, U, B\}$ representing government, non-Bridge private, non-enrollment, and attending Bridge at cost. ρ_j represents the sub-LATE for those who are induced to attend Bridge who would have otherwise been in educational environment j .

If a reader has a prior belief about the range of plausible values for some of the sub-LATEs, information from the first stage effects on each counterfactual and the overall LATE $\hat{\rho}$ may be informative about the effect of the remaining sub-LATEs. The weights $\hat{w}_j = \frac{\hat{\pi}_j}{\hat{\pi}_S}$ are observed, so that under assumptions for the effect of non-Bridge private schools, non-enrollment, and the effect of paying for Bridge (the direct effect of the scholarship), the relative effect of Bridge compared to public schools is given by

$$\rho_G = \frac{-\hat{\rho} - w_P \rho_P - w_U \rho_U - w_B \rho_B}{w_G} \quad (7)$$

The implied weights given the first stage are reported in Table 39. Figure 8 illustrates this for the primary grade sample and Figure 9 for the pre-primary grade sample.

The PAPSA discussed plausible ranges of these values given the published literature and we summarize that review here. The literature on the effects of private schools relative to public schools is mixed. A brief literature review suggests plausible estimates of private schools (relative to public

schools) ranging from positive (Bold et al. (2013) in Kenya, Dean and Jayachandran (2019) in India) to negative (Abdulkadiroğlu et al. (2018) in the US state of Louisiana) with several studies finding small or zero effects indicating no difference in the effects of private and public schools (Hsieh and Urquiola (2006) in Chile, and Muralidharan and Sundararaman (2015) in the Indian state of Andhra Pradesh).¹⁶ In considering the effects of non-enrollment relative to public school, it is important to note that most non-enrollment in this study comes from pupils in the youngest cohort who appear to have delayed school entry. Non-enrollment is rare in the primary school sample, so our literature reviewed focused on the effects of early childhood education. There is a large literature showing the effects of early childhood education on test scores in the short run (Dean and Jayachandran, 2019; Puma et al., 2012; Walters, 2015; Kline and Walters, 2016). There is some evidence on the effects of non-enrollment for older children. Burde and Linden (2013) find large positive effects on test scores from a program that increased access to local schools in Afghanistan. Finally, in considering the effect of the scholarship on those who would have paid to attend Bridge, we note that the literature on unconditional cash transfers usually finds relatively small effects on education (Haushofer and Shapiro, 2016), including test scores (Benhassine et al., 2015). There is also some evidence that the scholarship could have a negative effect if it changes the pupil’s perceived incentives to exert effort (Kremer et al., 2009), if parents pressure children less when education is free due to cognitive dissonance effects, or if teachers focus on paying students because they feel less accountable to the families of children who are on scholarship.

18 Non-experimental results on external validity

This section uses a non-experimental duration exposure identification strategy to estimate the effect of enrollment in Bridge using internal test score data supplied by Bridge. The purpose of this exercise is to compare the effect of Bridge on scholarship recipients to non-scholarship recipients to shed light on the external validity of the scholarship estimates. There are several reasons that scholarship recipients might gain more or less than other students. If Bridge focused resources on pupils who were on scholarship either because these students were further behind or in an effort to influence the results of the study, then the estimated effects, while internally valid, may not

¹⁶Muralidharan and Sundararaman (2015) did find positive effects on Hindi scores.

be reliable projections for many policy questions. Similarly, if Bridge teachers focus attention on paying customers over scholarship recipients or if students from families who are willing to pay for Bridge are those who stand to benefit the most from a Bridge education, the estimated effects on subject knowledge may understate the effect of Bridge for those families who choose to pay for it.

18.1 Data

Bridge provided data to the research team on enrollment and subject test scores for the universe of pupils enrolled in Bridge schools between 2016 and 2017, the two years of the scholarship program. Bridge also provided a file that identifies all pupils who attended Bridge through the scholarship program. This file was maintained for auditing purposes for the donor that provided funding for the scholarships.

Data on enrollment indicate the date at which the pupil first enrolled and, if applicable, the date at which the pupil withdrew.

Test data are a longitudinal panel of test scores on a variety of subjects related to the national curriculum, including religious studies. Bridge tests pupils up to seven times per year in each subject. At the beginning of the year there is a diagnostic test, followed by a mid-term and an end-term test in each of the three academic terms.

We focus on tests related to English language skills, Kiswahili language skills, and math. Because the data include multiple tests related to each of these subjects in the same time period, we refer to these as “subject groups”. English and Swahili have separate composition and reading tests as well as tests of grammar. In some grades, pupils have multiple tests related to one of these subjects. For example, some older students have separate tests for English composition, reading, and language, as well as a single English test. Kiswahili tests are separated along similar lines. To form a single test for each pupil, if multiple tests in a single subject group (English, Kiswahili, or math) are observed for the same pupil in the same time period, we form a single subject group test score by taking the average of standardized test scores for all scores in the same subject group and period.

In the analysis that follows we restrict to mid-term assessments only. To construct a balanced sample, we condition on pupils having all mid-term assessments for 2016 and 2017 (years 1 and 2). The purpose of restricting to a smaller number of assessments is because many pupils are missing at

least one assessment, limiting the sample size and reducing precision. By restricting to the sample with all mid-term assessments, we obtain a larger sample that is still balanced. Table 40 presents the relationship between endline scores and Bridge internal scores.

18.2 Identification strategy: duration exposure

The identification strategy in this section uses variation in the amount of time that a pupil has been enrolled at Bridge at the time of taking a particular test. First we show that new arrivals at Bridge tend to score below the average Bridge pupil and, over the course of a year or more, gradually converge to the mean. We view this pattern as suggestive non-experimental evidence of Bridge’s effect on new entrants.¹⁷

Figure 11 shows the average test score of pupils who entered Bridge in January or February 2016, a time period coinciding with the start of the scholarship program. The figure shows the averages separately for scholarship recipients and non-scholarship recipients.¹⁸ The figure shows that when pupils arrive at Bridge, they tend to score at least 0.2σ below the overall mean at Bridge schools.

These results motivate a test of the effect of the duration of exposure to Bridge. We assume that effect of increasing exposure to Bridge is linear in the amount of time that a pupil attends Bridge.

We estimate the following model of test scores as a function of the duration of exposure at time t , D_{it} :

$$y_{ist} = \beta_0 + \beta_1 D_{it} + \beta_2 Z_i + \beta_3 D_{it} \times Z_i + \lambda_i + \epsilon_{ist} \quad (8)$$

where y_{ist} is a score in subject s for pupil i at time t and $Z_i \in \{0, 1\}$ indicates that a pupil received a scholarship.

Equation 8 assumes that the effect of Bridge is linear. Figure 11 provides some visual evidence that the effects may plateau after longer periods of exposure as pupils converge to the mean. This would be expected if all pupils in Bridge schools are learning at a similar rate, but that new entrants must first catch up. We estimate Equation 8 using data from 200, 365 and 600 days.

¹⁷It should be noted that this pattern could also reflect adaptation to the Bridge environment, including learning how to take a test, without any increase in learning. The experimental results support the view that at least some of the upward trajectory in Bridge student test scores over time represents a causal effect of attendance.

¹⁸Note that some scholarship non-recipients may be participants in the study, but many are not in the study.

The assumption of linearity may be more plausible locally over smaller ranges of differences in exposure, although even for longer windows Figure 11 suggests that the linearity assumption may be approximately valid. We also estimate δ with and without pupil fixed effects.

Using the 200 day window, we find that a year of attendance at Bridge increases test scores by 0.588 standard deviations for non-scholarship recipients when pooling subject scores (Table 41). The interaction effect for scholarship recipients is -0.025 standard deviations and not statistically significant. Using the 365 day exposure window, we find that the effect of a year of Bridge attendance is 0.338 standard deviations. The interaction for scholarship recipients is 0.026 standard deviations and not statistically significant. Using the 600 day exposure window, the effect is 0.334 standard deviations for the non-recipients. The interaction for scholarship recipients is 0.036 standard deviations and not statistically significant.

In all specifications, we find that most of the effect of increasing exposure is experienced for both scholarship and non-scholarship students. We interpret this as evidence that the results of the experiment are likely to provide valid forecasts of the effects of Bridge on students who enroll in Bridge without scholarships.

19 Bridge academy staff characteristics

Bridge provided data on all teachers and academy managers in Bridge schools. Two files were provided: one showing staff employed as of August 2016 and the other showing staff employed as of February 2019. Data include gender, age, educational attainment, date of hire, and salary level. The data include teachers, academy managers, and substitute teachers. We report data on the cross section of staff who were employed as of 2016 and 2019 separately. Results are reported in Table 42.

20 Characteristics of Bridge pupils and classrooms

Bridge provided data on all pupils enrolled in Bridge schools during the time period of the scholarship program. We report results on a cross section of pupils enrolled in Bridge as of January 2017, the beginning of the second year of the scholarship program. Data include the academy location where the pupils is enrolled, gender, the start date of the pupil, and their grade. We compute

pupil-teacher ratios for each pupil as the total number of pupils in the same academy and grade. Nearly all Bridge schools had a single stream of students, so that the number of pupils in a grade captures the number of pupils in a classroom. Table 43 reports the results.

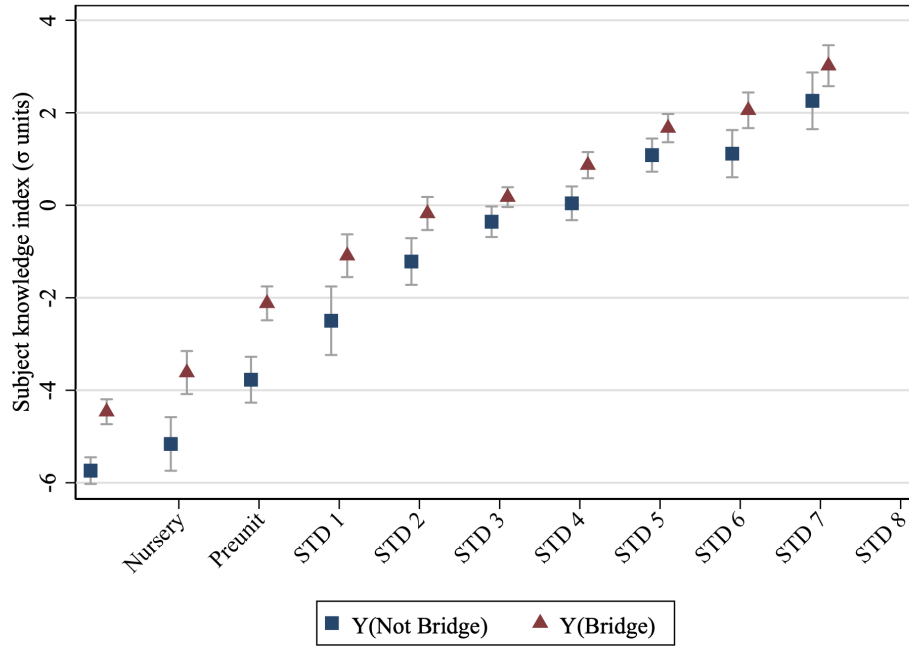
References

- ABADIE, A. (2002): “Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models,” *Journal of the American Statistical Association*, 97, 284–292.
- ABADIE, A., J. D. ANGRIST, AND G. IMBENS (2002): “Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings,” *Econometrica*, 70.
- ABADIE, A., M. M. CHINGOS, AND M. R. WEST (2018): “Endogenous Stratification in Randomized Experiments,” *The Review of Economics and Statistics*, 100.
- ABDULKADIROĞLU, A., P. A. PATHAK, AND C. R. WALTERS (2018): “Free to Choose: Can School Choice Reduce Student Achievement,” *American Economic Journal: Applied Economics*.
- ANGRIST, J. D. AND G. W. IMBENS (1995): “Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity,” *Journal of the American Statistical Association*, 90, 431–442.
- BENHASSINE, N., F. DEVOTO, E. DUFLO, P. DUPAS, AND V. POULIQUEN (2015): “Turning and Shove into a Nudge? A “Labeled Cash Transfer” for Education,” *American Economic Journal: Economic Policy*, 7, 86–125, [Link](#).
- BOLD, T., D. FILMER, G. MARTIN, E. MOLINA, C. ROCKMORE, B. STACY, J. SVENSSON, AND W. WANE (2017): “What do teachers know and do? Does it matter? Evidence from Primary schools in Africa,” .
- BOLD, T., M. KIMENYI, G. MWABU, AND J. SANDEFUR (2013): “The High Return to Private Schooling in a Low-Income Country,” Tech. rep., Africa Growth Initiative at Brookings, [Link](#).
- 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*.
- CHERNOZHUKOV, V. AND C. HANSEN (2013): “Quantile Models with Endogeneity,” *Annual Review of Economics*.
- CHETTY, R., J. N. FRIEDMAN, AND J. E. ROCKOFF (2014): “Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood,” *American Economic Review*, 104, 2633–79.
- DEAN, J. T. AND S. JAYACHANDRAN (2019): “Attending kindergarten improves cognitive but not socioemotional development in India,” Unpublished.
- GRAY-LOBE, G., A. KEATS, M. KREMER, I. MBITI, AND O. OZIER (2020): *Evaluation of Bridge International Academies in Kenya: Stage One Analysis and Plan for Subsequent Analysis*, AEA RCT Registry, [Link](#).
- HAUSHOFER, J. AND J. SHAPIRO (2016): “The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya,” *The Quarterly Journal of Economics*, [Link](#).
- HSIEH, C.-T. AND M. URQUIOLA (2006): “The effects of generalized school choice on achievement and stratification: Evidence from Chile’s voucher program,” *Journal of Public Economics*, 90, 1477–1503.

- KANE, T. J., J. E. ROCKOFF, AND D. O. STAIGER (2008): “What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City,” *Economics of Education Review*, 27, 615–631.
- KLINE, P. AND C. R. WALTERS (2016): “Evaluating Public Programs with Close Substitutes: the Case of Head Start,” *The Quarterly Journal of Economics*, 141, 1795–1848.
- KOLEN, M. J. AND R. L. BRENNAN (2014): *Test Equating, Scaling, and Linking*, New York, New York: Springer-Verlag.
- KREMER, M., E. MIGUEL, AND R. THORNTON (2009): “Incentives to Learn,” *Review of Economics and Statistics*.
- MURALIDHARAN, K. AND V. SUNDARARAMAN (2015): “The Aggregate Effect of School Choice: Evidence from a Two-stage Experiment in India,” *The Quarterly Journal of Economics*, 130, 1011–1066.
- PUMA, M., S. BELL, R. COOK, C. HEID, P. BROENE, F. JENKINS, A. MASHBURN, AND J. DOWNER (2012): “Third Grade Follow-up to the Head Start Impact Study Final Report,” Tech. rep., Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC, [Link](#).
- STAIGER, D. O. AND J. E. ROCKOFF (2010): “Searching for Effective Teachers with Imperfect Information,” *Journal of Economic Perspectives*, 24, 97–118.
- WALTERS, C. R. (2015): “Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start,” *American Economic Journal: Applied Economics*, 7, 76–102.
- WÜTHRICH, K. (2020): “A Comparison of Two Quantile Models with Endogeneity,” *Journal of Business & Economic Statistics*, 38, 443–456, [Link](#).

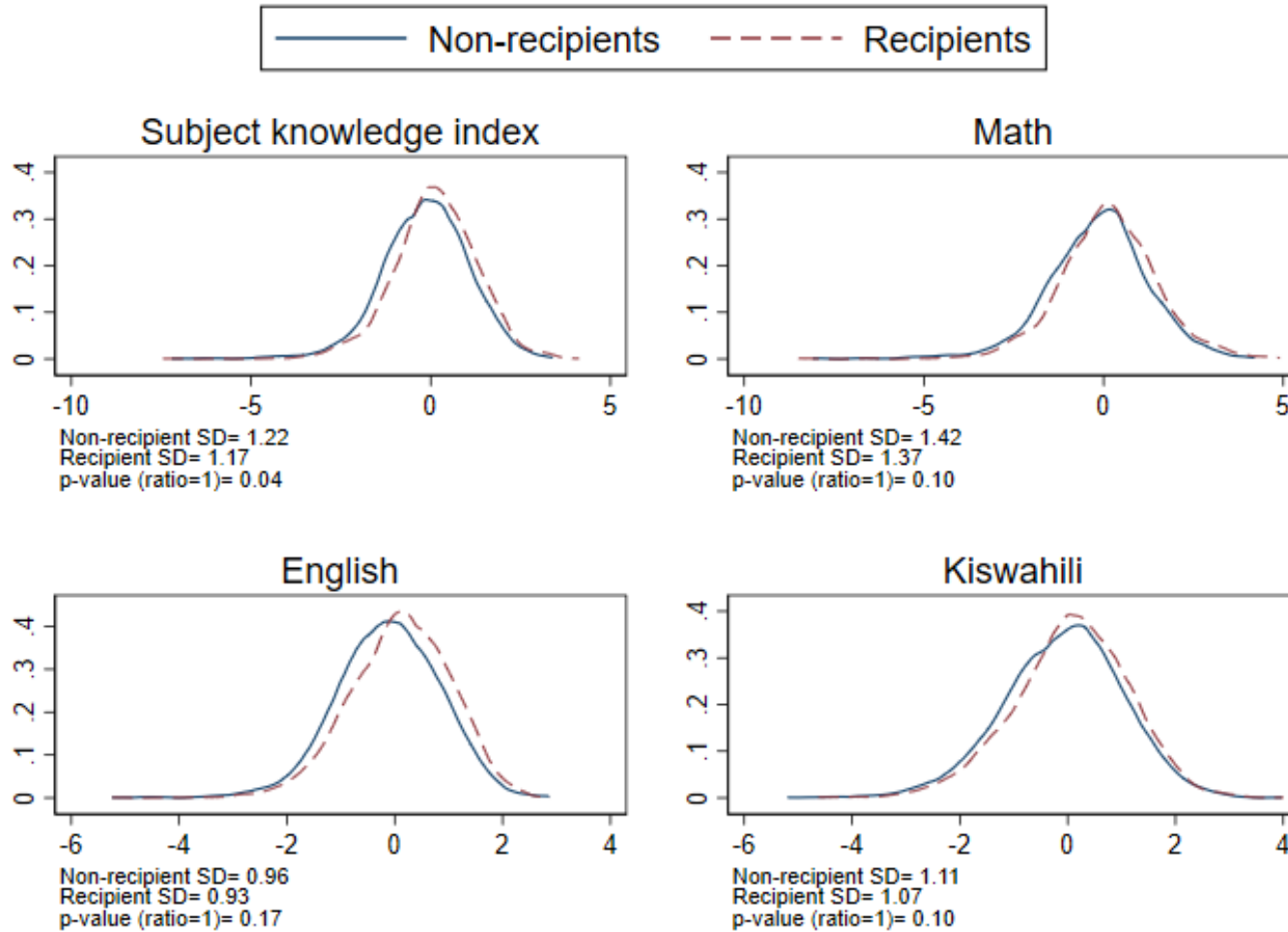
Figures

Figure 1: Complier mean potential outcomes by grade



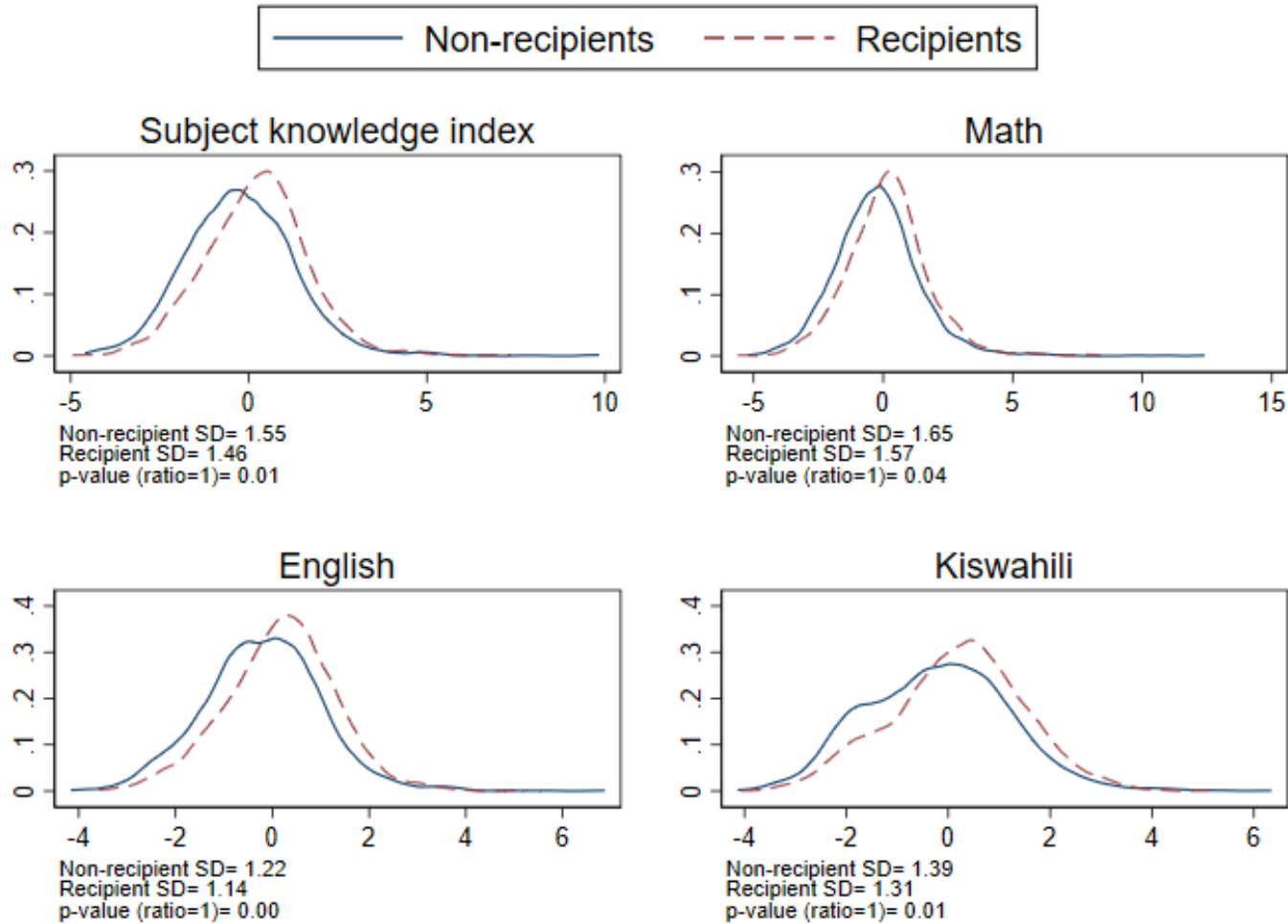
Notes: This figure graphs the estimated means of potential outcomes in Bridge, $Y(\text{Bridge})$, and not in Bridge, $Y(\text{Not Bridge})$, for compliers separately by projected year 2 (2017) grade using the approach of [Abadie \(2002\)](#). $Y(\text{Bridge})$ is the 2SLS estimate of attending Bridge $D_i \in \{0, 1\}$ on the transformed dependent variable $W_i^D = Y_i D_i$. $Y(\text{Not Bridge})$ is the 2SLS estimate of $1 - D_i$ on the transformed outcome $W_i^{1-D} = Y_i(1 - D_i)$ for $Y(\text{Bridge})$.

Figure 2: Distribution of Test scores by scholarship status: Primary school sample (P_{main})



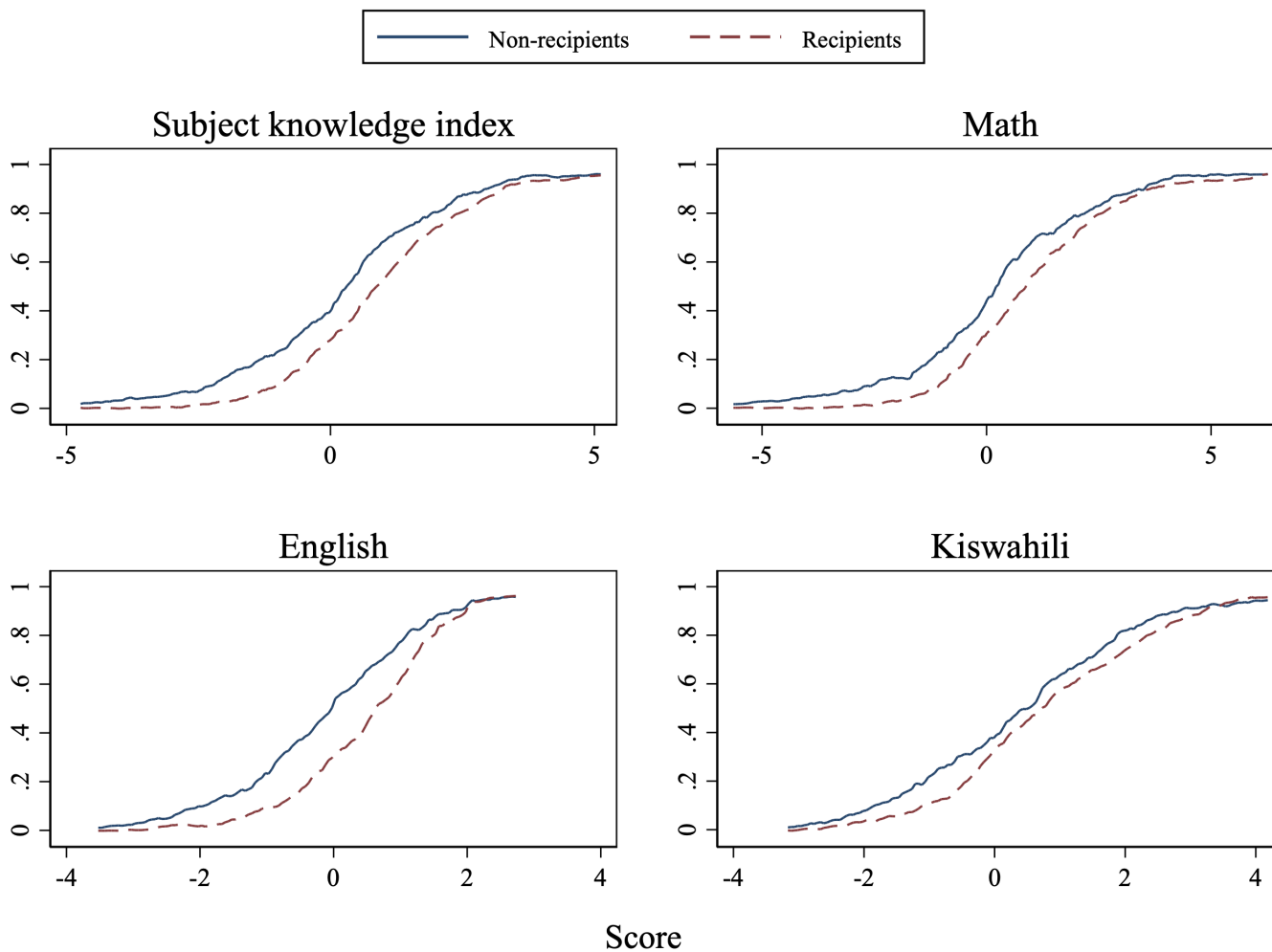
Notes: Figure shows the distribution of test scores for scholarship recipients and non-recipients in the P_{main} (Primary) sample. Test scores are residualized to account for baseline projected grade and strata fixed effects. Standard deviations for each group and a p-value from an F-test of the equality of the two standard deviations (that their ratio equals one) are reported below the figures.

Figure 3: Distribution of Test scores by scholarship status: Pre-primary sample (PP_{main})



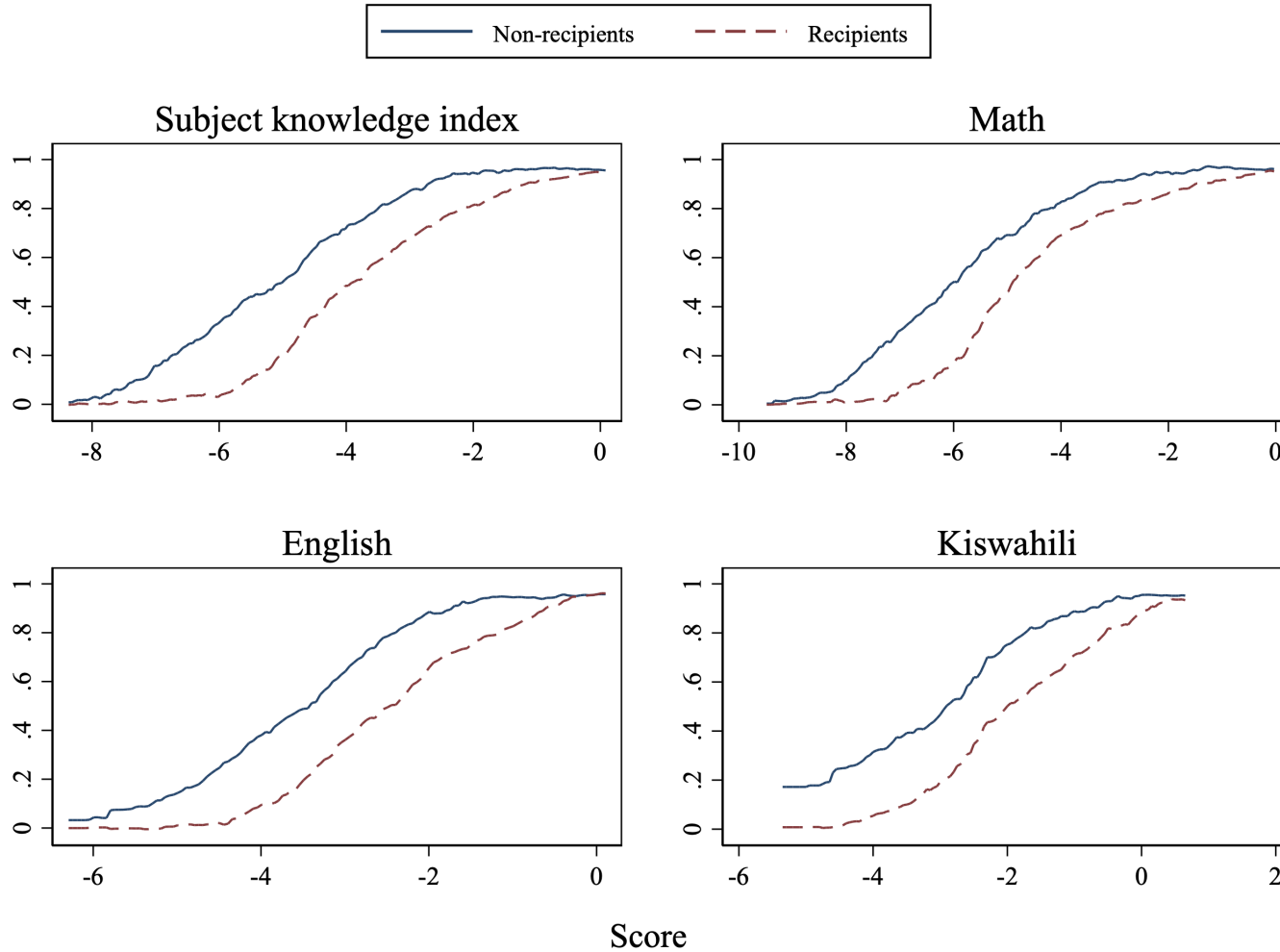
Notes: Figure shows the distribution of test scores for scholarship recipients and non-recipients in the PP_{main} (pre-primary) sample. Test scores are residualized to account for grade and strata fixed effects. Standard deviations for each group and a p-value from an F-test of the equality of the two standard deviations (that their ratio equals one) are reported below the figures.

Figure 4: Estimated distributions of outcomes of compliers: Primary school sample (P_{main})



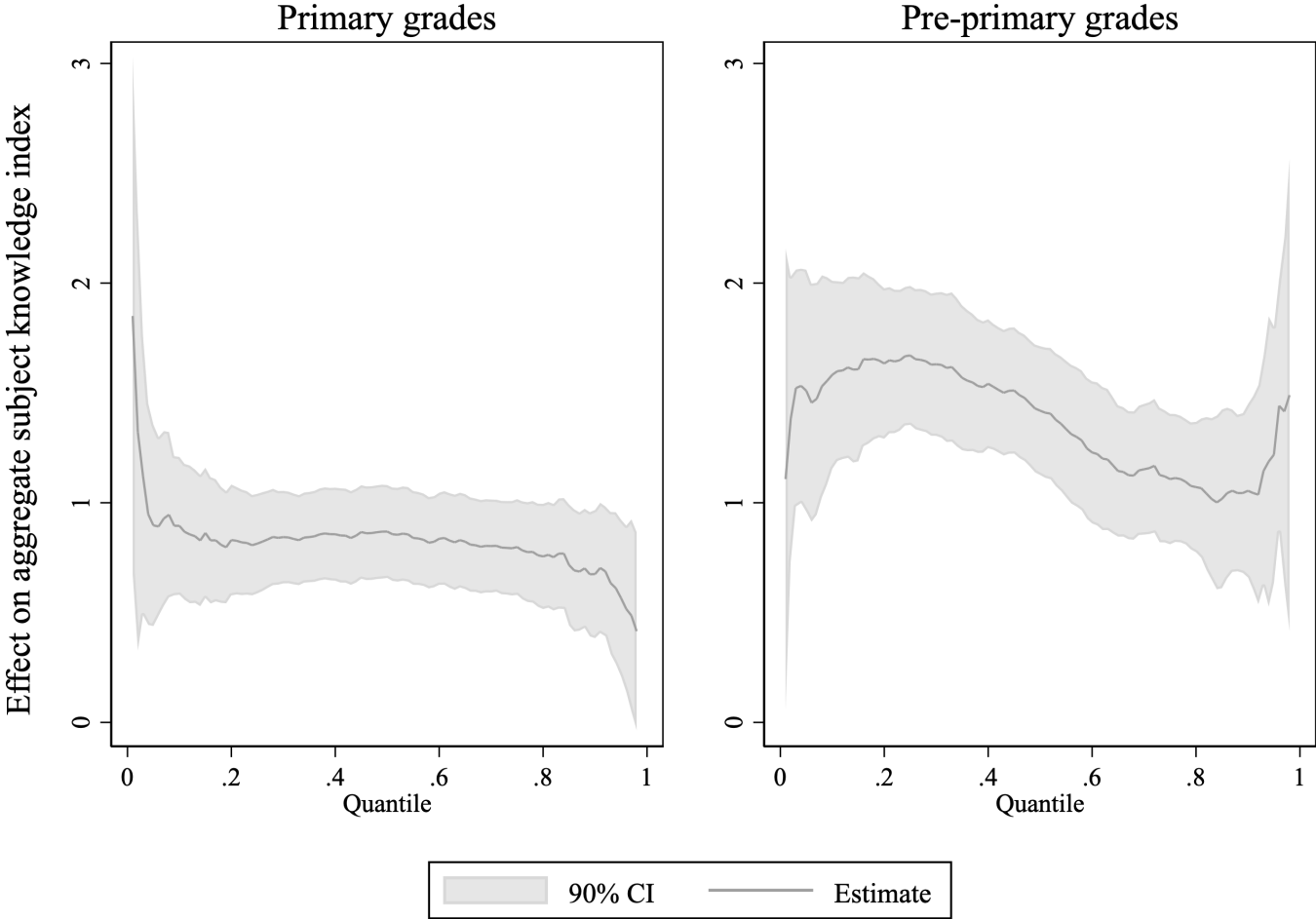
Notes: Figure shows the estimated distribution of test scores for scholarship recipients and non-recipients in the P_{main} (primary) sample using an approach from [Abadie \(2002\)](#). Each point in the estimated CDF reflects a separate 2SLS estimate. For recipient CDF gives the 2SLS estimated effect of attending Bridge $D_i \in \{0, 1\}$ on the transformed dependent variable $W_i^D = \mathbb{1}\{Y_i > x\}D_i$ where x represents a point in the support of the test score distribution. The non-recipient CDF gives the 2SLS estimated effect of not attending Bridge $1 - D_i$ on the transformed outcome $W_i^{1-D} = \mathbb{1}\{Y_i > x\}(1 - D_i)$.

Figure 5: Estimated distributions of outcomes of compliers: Pre-primary sample (PP_{main})



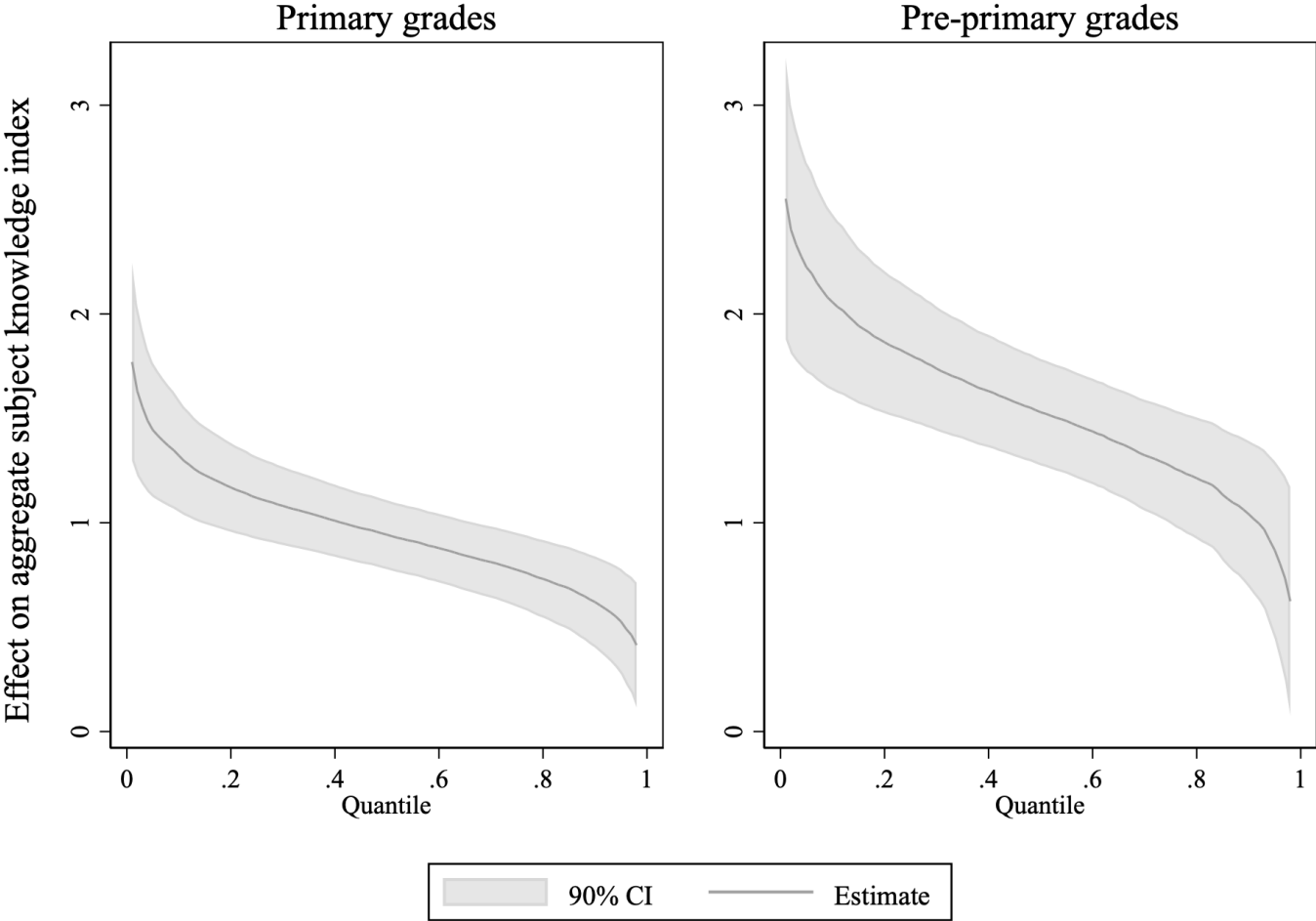
Notes: Figure shows the estimated distribution of test scores for scholarship recipients and non-recipients in the PP_{main} (pre-primary) sample using an approach from [Abadie \(2002\)](#). Each point in the estimated CDF reflects a separate 2SLS estimate. For recipient CDF gives the 2SLS estimated effect of attending Bridge $D_i \in \{0, 1\}$ on the transformed dependent variable $W_i^D = \mathbb{1}\{Y_i > x\}D_i$ where x represents a point in the support of the test score distribution. The non-recipient CDF gives the 2SLS estimated effect of not attending Bridge $1 - D_i$ on the transformed outcome $W_i^{1-D} = \mathbb{1}\{Y_i > x\}(1 - D_i)$.

Figure 6



Notes: Reports results from quantile regression using the procedure from [Abadie et al. \(2002\)](#).

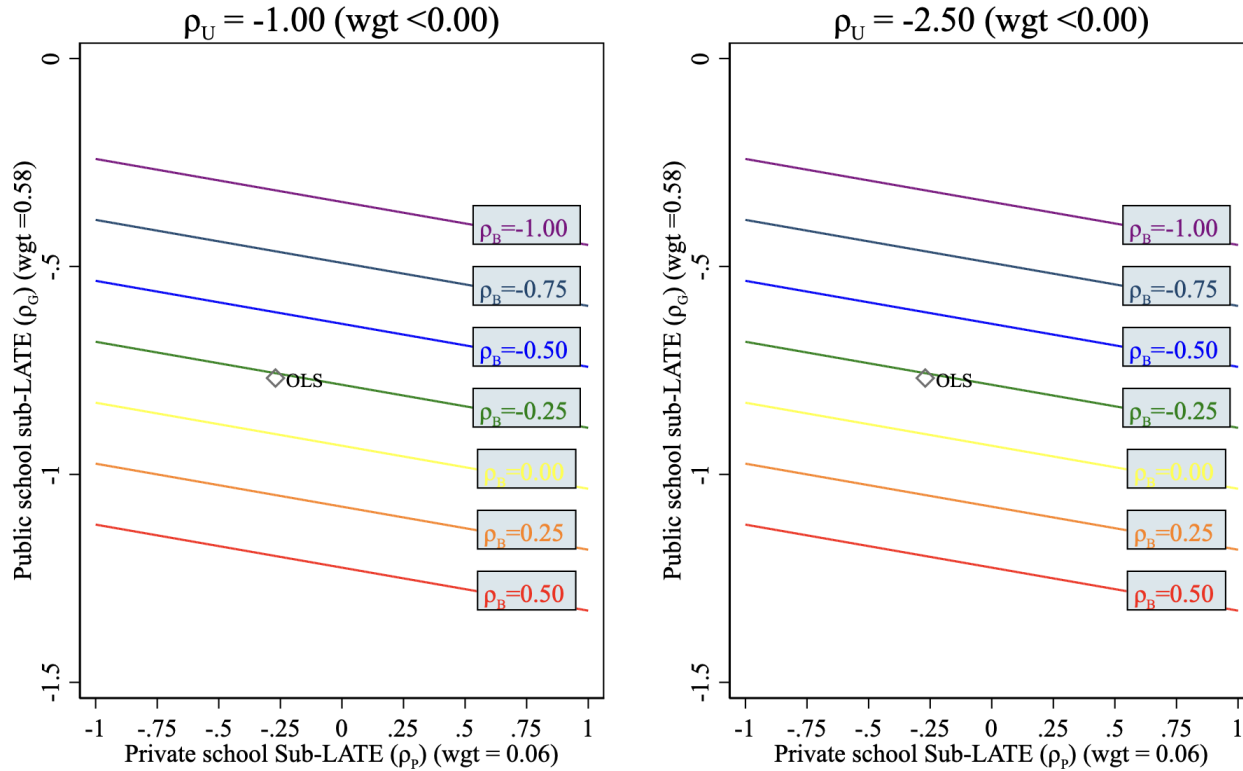
Figure 7



Notes: Reports results from quantile regression using the procedure from Chernozhukov and Hansen (2013).

Figure 8

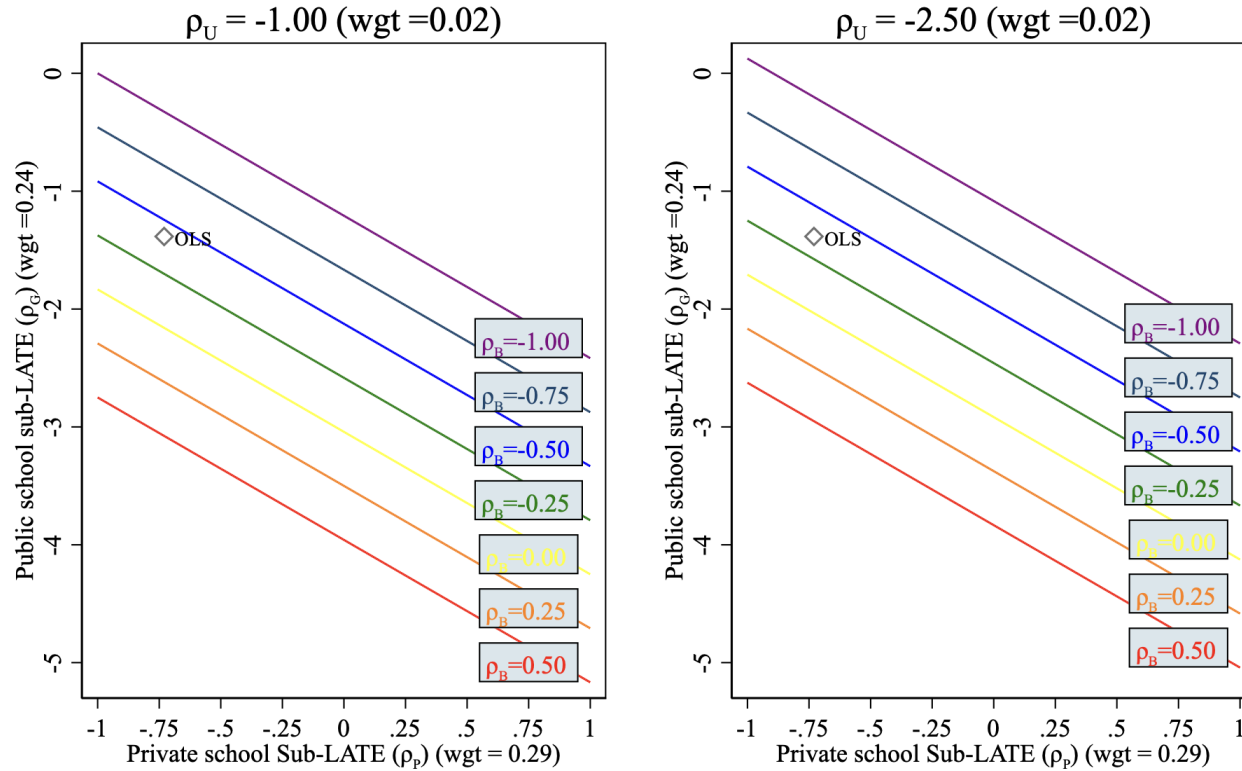
Contour plot of sub-LATEs using plug-in estimates Primary school sample



Notes: This figure illustrates the relationship between the separate sub-LATEs comparing Bridge to public (G), non-Bridge private (P), non-enrollment (U), and attending Bridge *at cost* (B) implied by the first stage effects on each of the counterfactuals. All reported effects reflect the difference between the potential outcome at Bridge for free and the potential outcome at the counterfactual. The weights implied by the first stage effect on attendance for public, non-Bridge private, non-enrollment, and the direct effect on always takers are 0.58, 0.06, 0.00, and 0.34. The figure also reports the implied sub-LATE comparing Bridge to public schools if OLS estimates of the effect of the remaining counterfactuals are assumed to be consistent estimates of the sub-LATEs. The OLS coefficients on attendance at public, non-Bridge private, non-enrollment, and attending Bridge without the scholarship are -0.55, -0.27, -2.53, and -0.23. The sample restricts to the P_{main} sample.

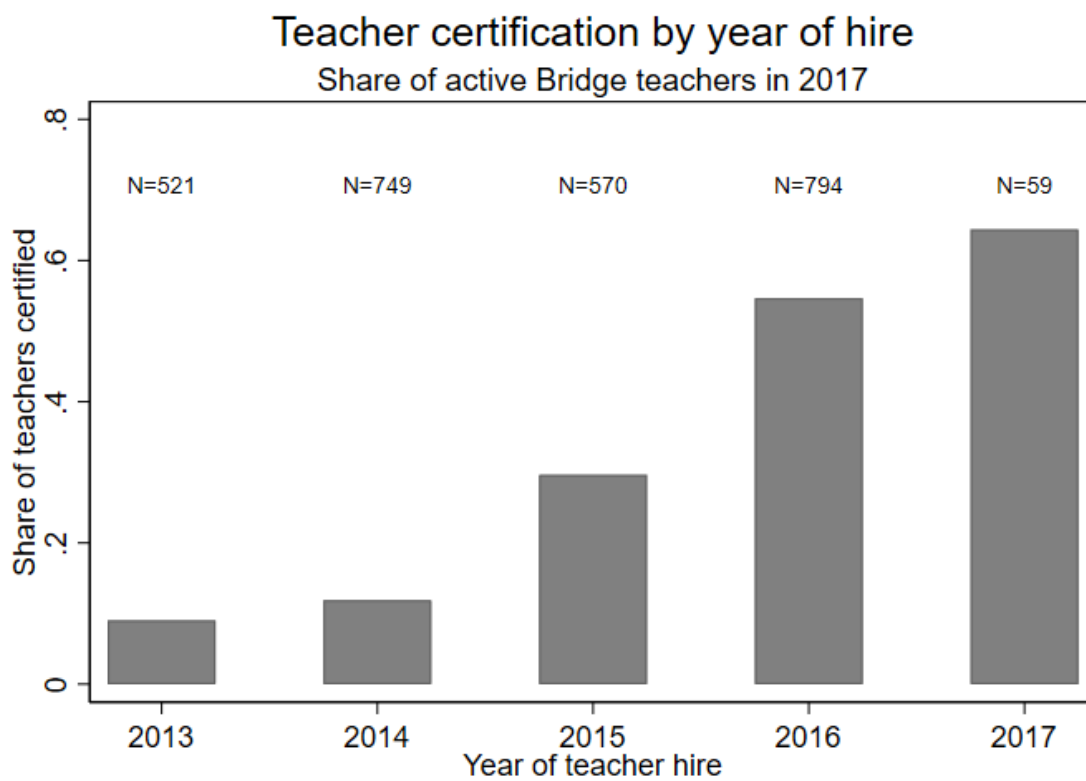
Figure 9

Contour plot of sub-LATEs using plug-in estimates Pre-primary sample



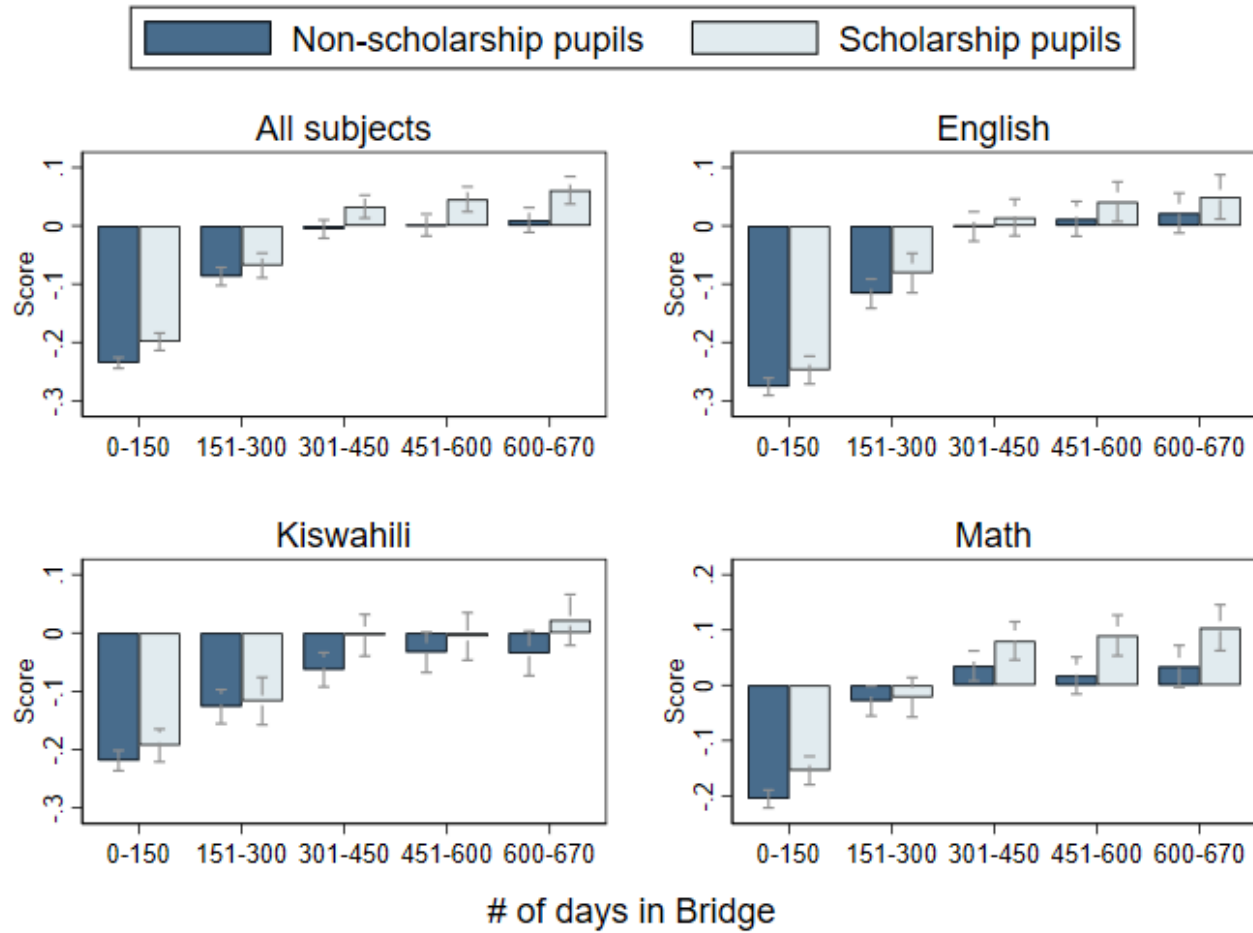
Notes: This figure illustrates the relationship between the separate sub-LATEs comparing Bridge to public (G), non-Bridge private (P), non-enrollment (U), and attending Bridge *at cost* (B) implied by the first stage effects on each of the counterfactuals. All reported effects reflect the difference between the potential outcome at Bridge for free and the potential outcome at the counterfactual. The weights implied by the first stage effect on attendance for public, non-Bridge private, non-enrollment, and the direct effect on always takers are 0.24, 0.29, 0.02, and 0.44. The figure also reports the implied sub-LATE comparing Bridge to public schools if OLS estimates of the effect of the remaining counterfactuals are assumed to be consistent estimates of the sub-LATEs. The OLS coefficients on attendance at public, non-Bridge private, non-enrollment, and attending Bridge without the scholarship are -0.88, -0.73, -3.04, and -0.33. The sample restricts to the PP_{main} sample.

Figure 10



Notes: Authors' calculations based on data provided by Bridge on new teacher hires. These results restrict to primary school level teachers.

Figure 11



Notes: Table shows the trajectory of standardized test scores for students in Bridge schools separately for the students who enrolled in Bridge on scholarship and those who did not.

Tables

Table 1: Estimated means of baseline covariates for non-recipient compliers

	Primary		Pre-primary	
	Full sample mean (1)	Non-recipient complier mean (2)	Full sample mean (3)	Non-recipient complier mean (4)
<i>Pupil characteristics</i>				
Female	0.51	0.45	0.49	0.44
Age	10.63	10.65	5.10	5.03
Mother is alive	0.91	0.90	0.94	0.94
Father is alive	0.74	0.76	0.80	0.81
Math percentile rank	0.68	0.63	0.68	0.71
English percentile rank	0.68	0.66	0.67	0.65
<i>Home characteristics</i>				
Floor of home is dirt	0.65	0.73	0.56	0.61
Walls of home are mud	0.55	0.63	0.47	0.47
Household has electricity	0.37	0.34	0.40	0.38
Household has latrine	0.73	0.78	0.72	0.71
HH income (annual 000s KES)	43.03	40.14	47.95	45.90
<i>Primary caregiver characteristics</i>				
Caregiver can read	0.12	0.14	0.10	0.13
Caregiver can write	0.15	0.15	0.13	0.17
<i>Caregiver occupation</i>				
Agriculture	0.16	0.19	0.10	0.14
Casual laborer	0.40	0.42	0.40	0.40
Small business owner	0.17	0.18	0.18	0.19

Notes: This table shows the estimated means of baseline covariates for scholarship non-recipient compliers alongside full sample means for primary and pre-primary groups. The non-recipient complier mean is estimated using the approach of [Abadie \(2002\)](#) and is the 2SLS estimate of $1 - D_i$ on the transformed outcome $V_i^{1-D} = X_i(1 - D_i)$. All specifications include linear control for probability of treatment assignment. For the primary school sample, the N is 4621 for all rows except *Mother is alive* (N = 3648), *Father is alive* (N = 3648), *Math percentile rank* (N = 4168) and *English percentile rank* (N = 4168). For the pre-primary school sample, the N is 3811 for all rows except *Mother is alive* (N = 3105), *Father is alive* (N = 3105), *Math percentile rank* (N = 1349) and *English percentile rank* (N = 1349).

Table 2: Intent-to-treat effect on stability of fixed applicant characteristics

	P _{main}		PP _{main}		P _{brig}		PP _{brig}		P _{priv}	
	Non-recipient mean (1)	Coef. (2)	Non-recipient mean (3)	Coef. (4)	Non-recipient mean (5)	Coef. (6)	Non-recipient mean (7)	Coef. (8)	Non-recipient mean (9)	Coef. (10)
Name string distance	0.84	-0.01 (0.05)	0.90	0.08 (0.07)	1.42	-0.08 (0.10)	1.57	-0.16 (0.22)	1.51	-0.28*** (0.12)
Observations	2,304	4,857	1,813	3,958	660	1,259	217	400	575	1,071
Endline age - baseline age	0.10	-0.03 (0.05)	0.08	-0.07 (0.07)	0.49	0.06 (0.08)	0.15	-0.13 (0.19)	0.44	-0.01 (0.10)
Observations	2,180	4,624	1,730	3,800	636	1,204	208	388	527	990
Endline age - baseline age	1.01	-0.04 (0.04)	1.24	-0.11** (0.05)	0.88	0.14** (0.07)	0.83	0.36** (0.16)	0.96	0.02 (0.07)
Observations	2,180	4,624	1,730	3,800	636	1,204	208	388	527	990
Pupil's gender matches	0.97	0.00 (0.00)	0.97	-0.01 (0.01)	0.98	-0.00 (0.01)	0.99	-0.02 (0.02)	0.96	0.01 (0.01)
Observations	2,181	4,625	1,730	3,801	636	1,204	208	388	527	990

Notes: Each row represents a separate specification. The prefix P refers to primary school, and the prefix PP refers to pre-primary. The subscript *brig* indicates the sample of pupils who were continuing Bridge students, and the P_{priv} sample is the sample of pupils who said they would attend private primary schools in the absence of the scholarship. To calculate the age difference we subtract 2017 minus their birth year reported in the application from the pupil's endline reported age in years. We test whether the scholarship affected this difference itself as well as the absolute difference. *Name distance* is the Levenshtein distance between the pupil's name in the application and the name reported at endline. All specifications include strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 3: Effects of scholarship and Bridge attendance on academic subject knowledge
IPA sub-sample only

	Primary			Pre-primary			EYS gradient (7)
	EYS		SD	EYS		SD	
	ITT (1)	2SLS (2)	2SLS (3)	ITT (4)	2SLS (5)	2SLS (6)	
<i>Panel A: Math & language scores</i>							
Subject knowledge index	0.42*** (0.08)	1.15*** (0.21)	1.05*** (0.19)	0.60*** (0.11)	1.54*** (0.29)	1.40*** (0.26)	0.91*** (0.02)
Higher order skills index	0.39*** (0.07)	1.07*** (0.18)	1.19*** (0.20)	0.42*** (0.10)	1.08*** (0.27)	1.20*** (0.30)	1.11*** (0.02)
Language knowledge index	0.43*** (0.10)	1.18*** (0.27)	0.75*** (0.17)	0.73*** (0.15)	1.89*** (0.39)	1.20*** (0.25)	0.64*** (0.02)
Math score	0.42*** (0.08)	1.14*** (0.22)	1.20*** (0.23)	0.51*** (0.10)	1.32*** (0.27)	1.39*** (0.28)	1.05*** (0.02)
English score	0.50*** (0.10)	1.37*** (0.28)	0.79*** (0.16)	0.75*** (0.15)	1.94*** (0.38)	1.13*** (0.22)	0.58*** (0.01)
Kiswahili score	0.36*** (0.11)	0.99*** (0.31)	0.60*** (0.19)	0.71*** (0.17)	1.84*** (0.43)	1.10*** (0.26)	0.60*** (0.02)
Observations		979			716		
<i>Panel B: Science & social studies</i>							
Science & social studies	0.45*** (0.14)	1.26*** (0.37)	0.64*** (0.19)	0.00 (0.11)	0.00 (0.25)	0.00 (0.13)	0.51*** (0.02)
Science	0.37*** (0.11)	1.02*** (0.31)	0.83*** (0.25)	-0.00 (0.11)	-0.01 (0.26)	-0.01 (0.21)	0.81*** (0.03)
Social studies	0.38*** (0.14)	1.06*** (0.38)	0.41*** (0.15)	-0.01 (0.10)	-0.02 (0.24)	-0.01 (0.09)	0.39*** (0.02)
Observations		970			195		
Local content score	0.31** (0.14)	0.85** (0.38)	0.38** (0.17)				0.45*** (0.03)
Observations		799					

Notes: Sample restricts to students that applied to the scholarship program through IPA. The standard deviation unit measures test score outcomes in standard deviations of the Standard 4 test score distribution. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. The higher order skills index restricts excludes subject test items that reflect remembering/recalling information. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 4: Effects of scholarship and Bridge attendance on academic subject knowledge
Entire sample including all observations from Meru and Nakuru

	Primary			Pre-primary			EYS gradient (7)
	EYS		SD	EYS		SD	
	ITT (1)	2SLS (2)	2SLS (3)	ITT (4)	2SLS (5)	2SLS (6)	
<i>Panel A: Math & language scores</i>							
Subject knowledge index	0.31*** (0.04)	0.80*** (0.10)	0.87*** (0.11)	0.51*** (0.05)	1.38*** (0.15)	1.50*** (0.16)	0.92*** (0.01)
Higher order skills index	0.28*** (0.04)	0.89*** (0.11)	0.80*** (0.10)	0.29*** (0.05)	0.95*** (0.16)	0.86*** (0.15)	1.11*** (0.01)
Language knowledge index	0.34*** (0.05)	0.63*** (0.09)	0.97*** (0.13)	0.65*** (0.07)	1.25*** (0.13)	1.93*** (0.20)	0.65*** (0.01)
Math score	0.29*** (0.04)	0.86*** (0.12)	0.81*** (0.11)	0.42*** (0.05)	1.30*** (0.16)	1.24*** (0.15)	1.06*** (0.01)
English score	0.43*** (0.05)	0.71*** (0.08)	1.22*** (0.14)	0.64*** (0.07)	1.12*** (0.12)	1.92*** (0.20)	0.59*** (0.01)
Kiswahili score	0.26*** (0.05)	0.45*** (0.09)	0.74*** (0.15)	0.65*** (0.07)	1.19*** (0.13)	1.94*** (0.22)	0.61*** (0.01)
Observations		5,208			4,135		
<i>Panel B: Science & social studies</i>							
Science & social studies	0.47*** (0.07)	0.71*** (0.10)	1.34*** (0.18)	0.10 (0.07)	0.16 (0.12)	0.31 (0.22)	0.53*** (0.01)
Science	0.33*** (0.06)	0.78*** (0.13)	0.94*** (0.16)	0.07 (0.07)	0.19 (0.18)	0.23 (0.21)	0.83*** (0.01)
Social studies	0.50*** (0.07)	0.56*** (0.07)	1.42*** (0.19)	0.08 (0.08)	0.10 (0.09)	0.24 (0.22)	0.39*** (0.01)
Observations		5,139			877		
Local content score	0.54*** (0.07)	0.60*** (0.08)	1.46*** (0.20)				0.41*** (0.01)
Observations		4,268					

Notes: This table reports the estimated effect of attending Bridge on subject test scores including all observations from Meru and Nakuru that were excluded from the final sample. As discussed in the text, some applications in Meru and Nakuru were dropped from the the analysis. This table shows that the results on test scores are robust to including these observations. The standard deviation unit measures test score outcomes in standard deviations of the Standard 4 test score distribution. The higher order skills index restricts excludes subject test items that reflect remembering/recalling information. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 5: Effects of scholarship and Bridge attendance on academic subject knowledge using different IRT equating methods

	Primary						Pre-primary					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Panel A: Math & language scores</i>												
Subject knowledge index	0.81*** (0.11)	0.79*** (0.10)	0.87*** (0.12)	0.93*** (0.12)	1.01*** (0.15)	0.43*** (0.07)	1.35*** (0.15)	1.11*** (0.12)	1.58*** (0.17)	1.61*** (0.18)	0.74*** (0.09)	0.95*** (0.09)
Higher order skills index	0.90*** (0.12)	0.87*** (0.11)	0.96*** (0.13)	1.00*** (0.13)	1.16*** (0.17)	0.46*** (0.07)	0.91*** (0.16)	0.68*** (0.13)	1.22*** (0.21)	1.17*** (0.21)	0.45*** (0.10)	0.71*** (0.09)
Language knowledge index	0.61*** (0.09)	0.61*** (0.09)	0.61*** (0.09)	0.62*** (0.09)	0.68*** (0.11)	0.39*** (0.07)	1.23*** (0.13)	1.12*** (0.12)	1.39*** (0.15)	1.42*** (0.15)	0.78*** (0.09)	0.88*** (0.09)
Math score	0.90*** (0.13)	0.84*** (0.12)	0.99*** (0.14)	1.10*** (0.16)	1.21*** (0.19)	0.40*** (0.08)	1.27*** (0.16)	0.93*** (0.12)	1.53*** (0.19)	1.56*** (0.19)	0.59*** (0.09)	0.91*** (0.09)
English score	0.68*** (0.08)	0.67*** (0.08)	0.72*** (0.09)	0.73*** (0.09)	0.68*** (0.09)	0.49*** (0.07)	1.11*** (0.12)	1.01*** (0.11)	1.31*** (0.14)	1.48*** (0.16)	0.66*** (0.07)	0.89*** (0.08)
Kiswahili score	0.45*** (0.10)	0.46*** (0.10)	0.42*** (0.09)	0.42*** (0.09)	0.58*** (0.13)	0.22*** (0.08)	1.18*** (0.14)	1.08*** (0.13)	1.27*** (0.14)	1.16*** (0.13)	0.78*** (0.10)	0.84*** (0.09)
Observations	4,587	4,587	4,587	4,587	4,587	4,587	3,805	3,805	3,805	3,805	3,805	3,805
<i>Panel B: Science & social studies</i>												
Science & social studies	0.72*** (0.10)	0.69*** (0.10)	0.80*** (0.12)	0.86*** (0.12)	1.24*** (0.20)	0.43*** (0.08)	0.15 (0.12)	0.16 (0.12)	0.04 (0.29)	0.04 (0.24)	0.01 (0.16)	0.81*** (0.24)
Science	0.77*** (0.14)	0.72*** (0.12)	0.74*** (0.13)	0.88*** (0.16)	1.38*** (0.32)	0.36*** (0.08)	0.13 (0.19)	0.13 (0.18)	0.10 (0.22)	0.10 (0.23)	0.30*** (0.00)	0.59*** (0.24)
Social studies	0.58*** (0.08)	0.58*** (0.08)	0.67*** (0.09)	0.61*** (0.08)	0.46*** (0.06)	0.50*** (0.08)	0.12 (0.09)	0.13 (0.09)	0.08 (0.17)	0.07 (0.13)	0.07 (0.06)	0.98*** (0.25)
Observations	4,526	4,526	4,526	4,526	4,526	4,526	794	794	794	794	794	794
Local content score	0.59*** (0.09)	0.59*** (0.08)	0.61*** (0.09)	0.59*** (0.08)	0.58*** (0.08)	0.55*** (0.08)						
Observations	3,750	3,750	3,750	3,750	3,750	3,750						
Method used for IRT equating												
Stocking-Lord	X						X					
Haebara		X						X				
Mean-mean			X						X			
Mean-gmean				X						X		
Mean-sigma					X						X	
Unequated raw score						X						X

Notes: This table reports the estimated 2SLS effects of attending Bridge using different IRT equating methods, as specified at the bottom of the table. The outcomes are measured in standard deviations of Standard 4 test score distribution. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effect. The higher order skills index restricts excludes subject test items that reflect remembering/recalling information. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 6: Effects of scholarship and Bridge attendance on academic subject knowledge using different standard deviation units

	Primary						Pre-primary					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Panel A: Math & language scores</i>												
Subject knowledge index	0.81*** (0.11)	0.60*** (0.08)	0.59*** (0.08)	0.43*** (0.06)	0.27*** (0.04)	0.43*** (0.07)	1.35*** (0.15)	0.81*** (0.09)	0.83*** (0.09)	0.67*** (0.07)	0.44*** (0.05)	0.95*** (0.09)
Higher order skills index	0.90*** (0.12)	0.64*** (0.08)	0.62*** (0.08)	0.42*** (0.06)	0.24*** (0.03)	0.46*** (0.07)	0.91*** (0.16)	0.50*** (0.09)	0.51*** (0.09)	0.39*** (0.07)	0.25*** (0.04)	0.71*** (0.09)
Language knowledge index	0.61*** (0.09)	0.57*** (0.08)	0.52*** (0.08)	0.40*** (0.06)	0.27*** (0.04)	0.39*** (0.07)	1.23*** (0.13)	0.83*** (0.09)	0.86*** (0.09)	0.73*** (0.08)	0.55*** (0.06)	0.88*** (0.09)
Math score	0.90*** (0.13)	0.56*** (0.08)	0.57*** (0.08)	0.41*** (0.06)	0.26*** (0.04)	0.40*** (0.08)	1.27*** (0.16)	0.72*** (0.09)	0.73*** (0.09)	0.58*** (0.07)	0.36*** (0.04)	0.91*** (0.09)
English score	0.68*** (0.08)	0.68*** (0.08)	0.61*** (0.08)	0.50*** (0.06)	0.33*** (0.04)	0.49*** (0.07)	1.11*** (0.12)	0.84*** (0.09)	0.86*** (0.09)	0.72*** (0.08)	0.53*** (0.06)	0.89*** (0.08)
Kiswahili score	0.45*** (0.10)	0.40*** (0.08)	0.36*** (0.08)	0.28*** (0.06)	0.21*** (0.05)	0.22*** (0.08)	1.18*** (0.14)	0.80*** (0.09)	0.82*** (0.09)	0.71*** (0.08)	0.54*** (0.06)	0.84*** (0.09)
Observations	4,587	3,962	4,587	4,587	4,587	4,587	3,805	3,805	3,805	3,805	3,805	3,805
<i>Panel B: Science & social studies</i>												
Science & social studies	0.72*** (0.10)	0.49*** (0.08)	0.50*** (0.08)	0.39*** (0.06)	0.39*** (0.06)	0.43*** (0.08)	0.15 (0.12)	0.37 (0.23)	0.40 (0.25)	0.29 (0.24)	0.08 (0.07)	0.81*** (0.24)
Science	0.77*** (0.14)	0.44*** (0.09)	0.42*** (0.08)	0.30*** (0.05)	0.29*** (0.05)	0.36*** (0.08)	0.13 (0.19)	0.21 (0.25)	0.22 (0.27)	0.16 (0.24)	0.05 (0.07)	0.59*** (0.24)
Social studies	0.58*** (0.08)	0.53*** (0.09)	0.58*** (0.08)	0.44*** (0.06)	0.44*** (0.06)	0.50*** (0.08)	0.12 (0.09)	0.35 (0.22)	0.38 (0.25)	0.30 (0.24)	0.09 (0.07)	0.98*** (0.25)
Observations	4,526	3,901	4,526	4,526	4,526	4,526	794	794	794	794	794	794
Local content score	0.59*** (0.09)	0.55*** (0.09)	0.59*** (0.09)	0.47*** (0.07)	0.47*** (0.07)	0.55*** (0.08)						
Observations	3,750	3,127	3,750	3,750	3,750	3,750						
Sample used for SD units												
Standard 4 (test takers)	X						X					
Projected grade (test takers)		X						X				
Projected grade			X						X			
Projected grade group (pre-/primary)				X						X		
Full sample (cross age)					X						X	
Raw IRT (unequated) score						X						X

Notes: This table reports the estimated 2SLS effects of attending Bridge using different standard deviation units. Columns 1 and 6 report the pre-specified units, standardizing the distribution of pupils taking the Standard 4 test form. Columns 2 and 7 report results standardizing scores within grades using the distribution of students who took the form of the test that the pupil was projected to receive at baseline. Columns 3 and 8 report results that standardize scores within baseline grades. The distribution of test scores within baseline grades is larger than that for test takers because it includes errors in the data and so includes greater variation in ages. Columns 4 and 9 report results that standardize scores within the pre-primary and primary samples separately. Columns 5 and 10 report results that standardize to the distribution of test scores for the full sample. All specifications include controls for baseline characteristics, planned 2016 grade, and strata dummies. The higher order skills index restricts subject test items that reflect remembering/recalling information. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 7: Effects of years of exposure to Bridge

	PP _{main} (1)	P _{main} (2)	PP _{brig} (3)	P _{brig} (4)	P _{priv} (5)
Subject knowledge index	0.751*** (0.082)	0.439*** (0.059)	-0.301 (0.583)	0.316* (0.176)	0.367* (0.194)
Language knowledge index	0.962*** (0.105)	0.468*** (0.073)	-0.887 (0.791)	0.307 (0.214)	0.427* (0.232)
Math score	0.618*** (0.075)	0.421*** (0.060)	0.069 (0.568)	0.321* (0.183)	0.329* (0.193)
English score	0.948*** (0.102)	0.578*** (0.075)	-0.719 (0.791)	0.383* (0.223)	0.545** (0.255)
Kiswahili score	0.975*** (0.111)	0.363*** (0.082)	-1.048 (0.835)	0.234 (0.254)	0.314 (0.243)
Number of observations	3,811	4,621	388	1,204	991

Notes: This table reports the results from 2SLS specifications where the endogenous variable representing the *number of years* a pupil enrolled at a Bridge school. Test score outcomes are measured in EYS units. All specifications include linear controls for the probability of treatment in the randomization strata. Robust standard errors are reported in parentheses.

Table 8: Effects of scholarship and Bridge attendance on specific items - Pre-primary only

	By projected 2017 grade											
	PP _{main}			Nursery			Pre-unit			Standard 1		
	Non-recipient			Non-recipient			Non-recipient			Non-recipient		
	mean	ITT	2SLS	mean	ITT	2SLS	mean	ITT	2SLS	mean	ITT	2SLS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
Letters per minute	15.80	4.15*** (0.55) 1,353	11.90*** (1.56) 2,938	13.45	4.60*** (0.66) 895	12.66*** (1.78) 1,974	19.94	3.40*** (1.21) 340	11.09*** (3.89) 718	21.73	2.58 (1.89) 118	7.28 (5.22) 246
<i>Pupil can...</i>												
...read the word <u>Thursday</u>	0.12	0.05*** (0.01) 1,723	0.14*** (0.03) 3,786	0.05	0.00 (0.01) 940	0.00 (0.03) 2,065	0.11	0.05* (0.03) 402	0.17** (0.08) 845	0.31	0.16*** (0.03) 381	0.44*** (0.10) 876
...read a simple English sentence	0.22	0.10*** (0.01) 1,723	0.30*** (0.04) 3,786	0.09	0.06*** (0.01) 940	0.16*** (0.04) 2,065	0.24	0.13*** (0.03) 402	0.44*** (0.12) 845	0.51	0.20*** (0.03) 381	0.55*** (0.10) 876
...read a simple Kiswahili sentence	0.11	0.06*** (0.01) 1,673	0.16*** (0.03) 3,683	0.04	0.01 (0.01) 932	0.02 (0.03) 2,048	0.09	0.09*** (0.02) 396	0.29*** (0.08) 832	0.32	0.15*** (0.04) 345	0.41*** (0.10) 803
...count eight stars	0.83	0.04*** (0.01) 1,235	0.13*** (0.04) 2,692	0.79	0.06*** (0.02) 895	0.16*** (0.05) 1,974	0.91	0.01 (0.02) 340	0.03 (0.07) 718			
...subtract 5 from 8 stars	0.42	0.05*** (0.02) 1,235	0.14*** (0.06) 2,692	0.37	0.04* (0.02) 895	0.11* (0.06) 1,974	0.58	0.07* (0.04) 340	0.21* (0.12) 718			
...add 4+2	0.50	0.07*** (0.01) 1,673	0.19*** (0.04) 3,683	0.34	0.06*** (0.02) 932	0.17*** (0.06) 2,048	0.62	0.08*** (0.03) 396	0.27*** (0.11) 832	0.80	0.06*** (0.03) 345	0.17*** (0.07) 803
...complete sequence 9,8,7,--	0.19	0.04*** (0.01) 1,673	0.12*** (0.04) 3,683	0.10	0.01 (0.01) 932	0.03 (0.04) 2,048	0.15	0.06** (0.03) 396	0.19** (0.09) 832	0.47	0.10*** (0.04) 345	0.28*** (0.10) 803

Notes: Table reports estimates of the effect of the scholarship and Bridge attendance on performance on specific items. *Letters per minute* indicates the number of correct letters identified in one minute. Pupils were shown 60 letters that included both upper and lower case letters. Some letters were repeated. Only 1.3 percent of pupils were able to identify 60 letters in 60 seconds. The simple English sentence was *Ken has a big dog*. The simple Kiswahili sentence was *Kaka alikunywa sharubati*. All specifications include controls for the probability of treatment for the pupil's randomization strata. Note that the predicted outcome for the ability to read a simple sentence in Standard 1 (Column 12) is inside the unit interval. Using the approach of [Abadie \(2002\)](#), we estimate the expected $Y(1)$ for compliers to be 0.82, and $Y(0)$ to be 0.27. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 9: Distributional effects of attending Bridge

	Quantile effects				Differences		FOSD test
	10 (1)	20 (2)	80 (3)	90 (4)	20-80 (5)	10-90 (6)	p-value (7)
Panel A: Primary grades							
Subject knowledge index	0.97*** (0.21)	0.90*** (0.17)	0.82*** (0.16)	0.74*** (0.19)	0.08 [0.08]	0.24*** [0.10]	0.004
Higher order skills index	0.90*** (0.19)	0.81*** (0.14)	0.67*** (0.13)	0.58*** (0.17)	0.14** [0.07]	0.32*** [0.12]	0.003
Language knowledge index	1.00*** (0.33)	1.21*** (0.25)	0.82*** (0.19)	0.75*** (0.20)	0.38*** [0.08]	0.24** [0.12]	0.014
Math	0.83*** (0.18)	0.81*** (0.16)	0.78*** (0.15)	0.77*** (0.20)	0.04 [0.08]	0.06 [0.12]	0.001
English	1.23*** (0.29)	1.25*** (0.24)	1.01*** (0.22)	0.88*** (0.22)	0.24*** [0.08]	0.36*** [0.09]	0.009
Kiswahili	0.78*** (0.31)	0.89*** (0.29)	0.75*** (0.21)	0.61*** (0.23)	0.14* [0.08]	0.17 [0.11]	0.006
Panel B: Pre-primary grades							
Subject knowledge index	1.72*** (0.28)	1.78*** (0.23)	1.17*** (0.20)	1.15*** (0.26)	0.61*** [0.11]	0.58*** [0.15]	0.008
Higher order skills index	0.97*** (0.20)	0.89*** (0.18)	0.69*** (0.21)	0.76*** (0.28)	0.20* [0.11]	0.21 [0.17]	0.005
Language knowledge index	2.55*** (0.38)	2.66*** (0.34)	1.38*** (0.24)	1.33*** (0.37)	1.28*** [0.10]	1.22*** [0.15]	0.025
Math	1.29*** (0.22)	1.28*** (0.20)	1.01*** (0.21)	1.10*** (0.31)	0.27*** [0.11]	0.19 [0.17]	0.007
English	2.64*** (0.42)	2.26*** (0.31)	1.47*** (0.25)	1.38*** (0.34)	0.79*** [0.09]	1.25*** [0.13]	0.011
Kiswahili	2.31*** (0.46)	3.10*** (0.37)	1.39*** (0.25)	1.26*** (0.41)	1.71*** [0.12]	1.05*** [0.18]	0.014

Notes: Quantile effects are from the method of [Abadie et al. \(2002\)](#). The specification for quantile regression includes planned 2016 grade, and control linearly for the randomization strata probability of scholarship assignment. Test score outcomes are expressed in units of equivalent years of schooling (EYS). Quantile analytical standard errors are reported in parentheses. Bootstrap standard errors of the difference in quantile effects are reported in brackets and represent that standard deviation of the statistic from 1000 bootstrap samples. The FOSD p-value reflects a test from [Abadie \(2002\)](#) of the null hypothesis that neither distribution is first order stochastically dominant (low p-values support first order stochastic dominance). The test is a modification of the Kolmogorov-Smirnov test of equality of distributions. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 10: Heterogeneous 2SLS effects of attending Bridge at cost on subject knowledge index
Individual characteristics

Covariate	Primary			Pre-primary			Pooled		
	Attended	× Cov	Cov	Attended	× Cov	Cov	Attended	× Cov	Cov
	Bridge			Bridge			Bridge		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Math percentile	1.11*** (0.36)	-0.45 (0.48)	0.90*** (0.20)	2.13*** (0.91)	-0.68 (1.19)	0.92* (0.54)	1.34*** (0.34)	-0.50 (0.45)	0.90*** (0.19)
		4,121			1,285			5,406	
English percentile	1.54*** (0.37)	-1.03** (0.50)	1.19*** (0.21)	2.88*** (0.96)	-1.86 (1.27)	1.41*** (0.56)	1.84*** (0.35)	-1.23*** (0.48)	1.25*** (0.20)
		4,121			1,284			5,405	
Self-regulation	1.46*** (0.33)	-0.42 (0.38)	0.85*** (0.14)	2.19*** (0.49)	0.52 (0.40)	0.31* (0.18)	1.53*** (0.24)	-0.09 (0.24)	0.64*** (0.11)
		932			673			1,605	
MDAT	2.86*** (1.04)	-0.12 (1.32)	0.83 (0.69)	1.99*** (0.37)	0.38 (0.37)	0.98*** (0.18)	2.06*** (0.34)	0.41 (0.35)	0.96*** (0.18)
		84			673			757	
Raven's score	1.25*** (0.30)	-0.32 (0.30)	0.71*** (0.11)				1.25*** (0.30)	-0.32 (0.30)	0.71*** (0.11)
		800						800	
LOO predicted index	0.84*** (0.14)	-0.51*** (0.19)	0.80*** (0.08)	1.69*** (0.21)	0.28 (0.35)	0.56*** (0.18)	1.25*** (0.12)	-0.18 (0.17)	0.70*** (0.08)
		4,587			3,805			8,392	

Notes: All specifications include randomization strata fixed effects. Test score units are equivalent years of schooling (EYS). English and math percentiles are the percentile rank of the pupil's self-reported Term 3 end-of-term test within grade-location/academy cells. The self-regulation, Malawi Development Assessment Tool (MDAT), and the Raven's score are standardized to have zero mean and unit standard deviation. The *Expected achievement index* is the predicted endline score from the leave-i-out regression of the aggregate learning index on a vector of baseline characteristics. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 11: Separate effects on males and females

	Primary				Pre-primary			
	OLS	2SLS		p-value	OLS	2SLS		p-value
	Female-male difference (1)	Male (2)	Female (3)		Female-male difference (5)	Male (6)	Female (7)	
Subject knowledge index (EYS)	0.24*** (0.06) 2,180	1.03*** (0.18) 2,284	0.75*** (0.16) 2,337	0.25	0.35*** (0.08) 1,735	1.69*** (0.22) 1,949	1.37*** (0.26) 1,862	0.35
Subject knowledge index (EYS)	0.45*** (0.08) 2,180	1.21*** (0.22) 2,284	0.66*** (0.19) 2,337	0.06	0.54*** (0.11) 1,735	2.24*** (0.28) 1,949	1.65*** (0.34) 1,862	0.18
Math score (EYS)	0.12* (0.06) 2,180	0.91*** (0.18) 2,284	0.80*** (0.16) 2,337	0.67	0.23*** (0.08) 1,735	1.34*** (0.20) 1,949	1.18*** (0.25) 1,862	0.62
Raven's matrices	-0.16*** (0.04) 2,181	0.25* (0.13) 2,288	-0.02 (0.12) 2,337	0.13	0.07** (0.03) 1,730	0.14 (0.09) 1,940	0.05 (0.11) 1,861	0.54
Digit span recall	0.05 (0.04) 2,152	0.14 (0.11) 2,262	0.32*** (0.12) 2,313	0.25	0.08* (0.04) 1,564	0.46*** (0.11) 1,742	0.09 (0.13) 1,718	0.03
Head-knees task					0.05 (0.06) 1,330	0.17 (0.15) 1,499	0.28 (0.20) 1,393	0.65
Unusual uses	-0.05 (0.05) 1,750	0.08 (0.13) 1,812	-0.01 (0.12) 1,870	0.61	-0.11*** (0.04) 1,335	-0.04 (0.12) 1,472	-0.03 (0.14) 1,432	0.95
Reached projected grade	0.04* (0.02) 2,180	0.24*** (0.05) 2,286	0.18*** (0.05) 2,337	0.37	0.02 (0.02) 1,724	0.18*** (0.05) 1,927	0.20*** (0.06) 1,856	0.90
Took KCPE	0.00 (0.03) 747	0.10 (0.09) 774	0.20*** (0.08) 842	0.38				
KCPE score	-3.68 (4.60) 577	30.29** (13.90) 611	6.80 (11.24) 682	0.19				
Secondary school transition	-0.03 (0.04) 527	-0.05 (0.11) 518	0.09 (0.10) 559	0.36				

Notes: All specifications include randomization strata fixed effects. Test score units are equivalent years of schooling (EYS). English and math percentiles are the percentile rank of the pupil's self-reported Term 3 end-of-term test within grade-location/academy cells. The self-regulation, Malawi Development Assessment Tool (MDAT), and the Raven's score are standardized to have zero mean and unit standard deviation. The *Expected achievement index* is the predicted endline score from the leave-i-out regression of the aggregate learning index on a vector of baseline characteristics. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 12: Effects of Bridge teacher/classroom characteristics

Covariate	2SLS with interaction terms						Pupil FE test	
	Primary			Pre-primary			Enrolled at Bridge	Not enrolled at Bridge
	Attended			Attended			Cov	Cov
	Bridge	× Cov	Cov	Bridge	× Cov	Cov	(7)	(8)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First year teaching at Bridge								
Coefficient	0.82***	-1.08	0.21	2.04***	0.30	-0.97	0.04	-0.01
SE	(0.12)	(0.70)	(0.18)	(0.36)	(2.40)	(0.82)	(0.12)	(0.08)
N tests		13,101			1,853		5,655	9,259
N pupils		4,473			618		3,550	4,319
Covariate mean		0.02			0.05		0.02	0.03
Covariate SD		0.15			0.21		0.14	0.18
More than 2 years experience								
Coefficient	0.65***	0.21	-0.10	3.76***	-2.21	1.04*	0.00	-0.03
SE	(0.19)	(0.20)	(0.08)	(1.39)	(1.52)	(0.60)	(0.03)	(0.03)
N tests		13,101			1,853		5,655	9,259
N pupils		4,473			618		3,550	4,319
Covariate mean		0.72			0.73		0.71	0.72
Covariate SD		0.45			0.44		0.45	0.45
Years experience teaching at Bridge								
Coefficient	0.50**	0.12*	-0.05*	3.40***	-0.55*	0.23	-0.00	-0.01
SE	(0.22)	(0.08)	(0.03)	(1.03)	(0.33)	(0.15)	(0.01)	(0.01)
N tests		13,101			1,853		5,655	9,259
N pupils		4,473			618		3,550	4,319
Covariate mean		2.42			2.41		2.44	2.41
Covariate SD		1.22			1.27		1.24	1.22
Certified teacher (TSC)								
Coefficient	0.89***	-0.25	0.03	1.88***	0.44	-0.04	-0.07**	-0.02
SE	(0.14)	(0.18)	(0.07)	(0.45)	(1.04)	(0.43)	(0.03)	(0.03)
N tests		13,101			1,853		5,655	9,259
N pupils		4,473			618		3,550	4,319
Covariate mean		0.32			0.28		0.32	0.31
Covariate SD		0.47			0.45		0.47	0.46
KCSE score (standardized)								
Coefficient	0.74***	-0.08	0.04	1.97***	0.26	0.06	-0.01	0.03*
SE	(0.13)	(0.10)	(0.04)	(0.44)	(0.35)	(0.17)	(0.02)	(0.02)
N tests		8,693			1,282		3,435	5,812
N pupils		3,821			430		2,228	2,728
Covariate mean		0.02			-0.13		-0.02	0.01
Covariate SD		0.96			1.00		0.98	0.96

Notes: This table reports results from stacking English, Kiswahili, and math test scores. Test score units are equivalent years of schooling (EYS). Aside from teacher/classroom value-added, teacher characteristics are only observed for the oldest cohort of the pre-primary sample. All specifications include controls for expected 2016 grade level, subject, and the probability of scholarship assignment for the randomization strata. Column 7 reports estimates from a pupil fixed effect model of the effect of the characteristic on subject test scores for all grade levels for which there is variation in the teacher characteristic. Column 7 reports results for pupils who were enrolled in Bridge in 2017. Reported coefficient is estimated by instrumenting the Bridge classroom value-added corresponding to the actual grade achieved in 2017 with the Bridge classroom value-added of the pupil's projected 2017 grade. In Standard 3 and below, Bridge has classroom teachers, so there is no variation in teacher characteristics that do not vary at the subject level (experience, certification). Column 8 reports analogous results from a placebo test for pupils who were not enrolled in Bridge in 2017. For tests of heterogeneous effects of attending Bridge, standard errors are clustered at the pupil level. 2SLS specifications include controls for planned 2016 grade, subject and the probability of treatment in the strata cell. For pupil FE estimates, the specifications also include subject dummies and standard errors are clustered at the academy-grade level. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 13: Correlation matrix of teacher characteristics

	Certified teacher (1)	KCSE score (2)	Years experience (3)
Certified teacher	1.000		
KCSE score	0.003	1.000	
Years of experience	-0.423***	-0.113***	1.000

Notes: Table reports the correlation matrix for teacher characteristics at Bridge schools. This table shows results for all teachers Teacher experience is the number of years the teacher has been teaching at Bridge. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 14: Heterogeneous 2SLS effects by class size

Outcome	Primary		Pre-primary		Pooled	
	Attended Bridge (1)	× Class size (2)	Attended Bridge (3)	× Class size (4)	Attended Bridge (5)	× Class size (6)
<i>Panel A: Intermediate outcomes</i>						
Number textbooks	0.830** (0.363)	-0.013 (0.020)	0.295 (0.451)	0.009 (0.023)	0.586** (0.282)	-0.003 (0.015)
	3,536		2,714		6,250	
Classroom crowding index	-0.379** (0.182)	0.001 (0.010)	0.000 (0.239)	-0.009 (0.012)	-0.225 (0.145)	-0.003 (0.008)
	3,541		2,927		6,468	
Teacher-classroom engagement index	0.279 (0.187)	-0.011 (0.011)	0.119 (0.577)	0.007 (0.023)	0.251 (0.177)	-0.007 (0.009)
	3,506		734		4,240	
Parental engagement index	0.489*** (0.185)	-0.009 (0.012)	0.686*** (0.241)	0.006 (0.014)	0.531*** (0.148)	0.001 (0.009)
	3,437		2,825		6,262	
Instructional adaptation index	0.459*** (0.148)	-0.007 (0.008)	0.187 (0.167)	0.004 (0.009)	0.351*** (0.110)	-0.003 (0.006)
	3,556		2,952		6,508	
<i>Panel B: Academic subject knowledge</i>						
Subject knowledge index (EYS)	0.979*** (0.281)	-0.008 (0.016)	1.348*** (0.368)	-0.000 (0.019)	1.081*** (0.226)	0.000 (0.012)
	3,567		2,975		6,542	
Subject knowledge index (EYS)	1.117*** (0.334)	-0.012 (0.019)	1.592*** (0.486)	0.006 (0.025)	1.240*** (0.283)	0.003 (0.015)
	3,567		2,975		6,542	

Notes: Class size is calculated using Bridge’s administrative test score data. The number of pupils in a class is calculated as the number of pupils with test scores on the 2017 Term 1. Specifications include indicators for randomization strata. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 15: Heterogeneous 2SLS effects of attending Bridge on subject knowledge index
Academy level characteristics

Covariate	Primary		Pre-primary		Pooled	
	Attended Bridge (1)	× Covariate (2)	Attended Bridge (3)	× Covariate (4)	Attended Bridge (5)	× Covariate (6)
Mean county KCPE score (std)	0.86*** (0.13)	0.10 (0.14)	1.47*** (0.16)	0.10 (0.18)	1.15*** (0.10)	0.06 (0.11)
		4,526		3,758		8,284
Median KCPE nearest govt sch	0.87*** (0.12)	-0.02 (0.10)	1.49*** (0.16)	0.08 (0.16)	1.15*** (0.10)	-0.00 (0.09)
		4,621		3,811		8,432
Yrs of Bridge acad operation	0.31 (0.27)	0.23** (0.11)	1.80*** (0.37)	-0.11 (0.12)	0.90*** (0.22)	0.10 (0.08)
		4,621		3,811		8,432
Rural location	1.16*** (0.20)	-0.46** (0.23)	1.32*** (0.22)	0.32 (0.32)	1.28*** (0.15)	-0.21 (0.19)
		4,621		3,811		8,432
Urban location	0.76*** (0.13)	0.48* (0.28)	1.56*** (0.20)	-0.22 (0.33)	1.08*** (0.11)	0.23 (0.21)
		4,621		3,811		8,432
Peri-urban location	0.85*** (0.12)	0.16 (0.33)	1.52*** (0.17)	-0.27 (0.45)	1.14*** (0.10)	0.03 (0.28)
		4,621		3,811		8,432
IHS HH monthly income	0.86*** (0.12)	-0.13 (0.12)	1.48*** (0.16)	0.05 (0.14)	1.15*** (0.10)	-0.05 (0.09)
		4,621		3,811		8,432

Notes: All specifications include baseline controls and a linear control for the probability of scholarship assignment in the applicants randomization stratum. Outcome test scores are expressed in equivalent years of schooling (EYS). Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 16: Sargan test for heterogeneous academy effects

	Primary grades (1)	Pre-primary grades (2)	Pooled (3)
Subject knowledge index			
Sargan	176.5	153.1	243.6
DF	177	156	277
p-value	0.496	0.550	0.926
N pupils	3,803	2,815	7,932
N Academies	180	158	279
Language knowledge index			
Sargan	162.3	154.2	229.6
DF	177	156	277
p-value	0.779	0.525	0.983
N pupils	3,803	2,815	7,932
N Academies	180	158	279
Math score			
Sargan	181.0	154.3	258.9
DF	177	156	277
p-value	0.404	0.523	0.775
N pupils	3,803	2,815	7,932
N Academies	180	158	279
English score			
Sargan	156.7	151.7	228.9
DF	177	156	277
p-value	0.861	0.581	0.984
N pupils	3,803	2,815	7,932
N Academies	180	158	279
Kiswahili score			
Sargan	174.4	156.7	239.6
DF	177	156	277
p-value	0.541	0.469	0.949
N pupils	3,803	2,815	7,932
N Academies	180	158	279

Notes: This table reports results from a Sargan over-identification test using interaction of treatment with location (academy) indicators. All specifications include controls for the baseline projected grade in 2016, academy dummies and the probability of treatment for the strata. All tests restrict to locations with 10 or more observations.

Table 17: Effects of scholarship and Bridge enrollment on spoon task

	Primary			Pre-primary			EYS gradient (7)
	EYS		SD	EYS		SD	
	ITT	2SLS	2SLS	ITT	2SLS	2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)	
<i>Spoons task</i>							
Unusual uses	0.63 (0.86)	1.84 (2.48)	0.07 (0.09)	-0.33 (0.81)	-1.00 (2.46)	-0.04 (0.09)	0.04*** (0.01)
		3,615			2,887		
Total uses	-0.02 (0.37)	-0.05 (1.07)	-0.00 (0.09)	-0.14 (0.38)	-0.44 (1.16)	-0.04 (0.10)	0.08*** (0.01)
		3,615			2,887		
Unusual uses (external judge)	0.24 (0.64)	0.71 (1.86)	0.04 (0.10)	-0.31 (0.58)	-0.95 (1.77)	-0.05 (0.09)	0.05*** (0.01)
		3,615			2,887		

Notes: This table shows the effect of the scholarship and Bridge enrollment on outcomes related to the spoon use task. The spoon use task asked children to enumerate as many uses of a spoon as they could. During the interview field officers classified individual by indicated if a response was unusual. The task is then scored by summing the number of responses (fluency) and the number of unusual responses (originality). To evaluate the sensitivity of the scoring to subjective field officer judgments, a later activity had external judges on mTurk classify responses as either unoriginal, somewhat original, or very original. Each item was rates by over 240 individual judges. At the item level, we calculate the fraction of judges who assessed the response to belong to each category. We then calculate an analogous originality score by summing the shares saying that the response was either somewhat or very original. All specifications include controls for baseline characteristics, planned 2016 grade, and strata dummies. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 18: Effects on occupational aspirations

	Primary grades			Pre-primary			Pooled		
	Non-recipient			Non-recipient			Non-recipient		
	mean	ITT	2SLS	mean	ITT	2SLS	mean	ITT	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Occupation academic achievement index	-0.019	0.010** (0.005)	0.029** (0.014)	-0.186	-0.001 (0.009)	-0.002 (0.026)	-0.092	0.005 (0.005)	0.015 (0.014)
Female occupation index	0.499	-0.001 (0.006)	-0.002 (0.018)	0.480	0.004 (0.007)	0.012 (0.021)	0.491	0.002 (0.005)	0.005 (0.014)
Aspires to professional occupation	0.626	0.017 (0.014)	0.049 (0.039)	0.398	-0.013 (0.016)	-0.038 (0.046)	0.526	0.004 (0.011)	0.010 (0.030)
<i>Specific occupations</i>									
Doctor	0.257	0.001 (0.013)	0.002 (0.037)	0.157	0.001 (0.012)	0.003 (0.035)	0.214	0.001 (0.009)	0.004 (0.026)
Teacher	0.120	-0.009 (0.010)	-0.025 (0.027)	0.158	-0.007 (0.012)	-0.020 (0.035)	0.136	-0.008 (0.008)	-0.023 (0.022)
Engineer	0.104	0.011 (0.009)	0.032 (0.026)	0.022	-0.010** (0.004)	-0.028** (0.013)	0.069	0.001 (0.005)	0.002 (0.015)
Pilot	0.159	-0.004 (0.011)	-0.010 (0.030)	0.138	0.008 (0.012)	0.024 (0.034)	0.150	0.001 (0.008)	0.004 (0.023)
Driver	0.052	-0.005 (0.006)	-0.013 (0.018)	0.095	-0.002 (0.010)	-0.005 (0.028)	0.071	-0.003 (0.006)	-0.008 (0.016)
Joint F-test		0.975			1.127		1.233		
p-value		0.521			0.255		0.126		
Observations		8,270			3,708		4,562		

Notes: Table reports the effects of the scholarship and attending Bridge on occupational aspirations. *Occ academic achievement index* is the expected math score given the occupation to which the pupil aspires. The prediction is formed on the control group only, and all predictions leave out individual i . *Female occupation index* is the probability that a pupil is female given the occupation to which they aspire. Predictions are formed on the control and leave i out in all cases. All responses are pupil reported. Robust standard errors are reported in parentheses. An F-statistic and corresponding p-value of the joint test of independence of treatment and dummies for each occupation is presented at the bottom of the table. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 19: Effects on social attitudes

	Primary grades			Pre-primary		
	Non-recipient			Non-recipient		
	mean	ITT	2SLS	mean	ITT	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Girls and boys pooled</i>						
Trusts people in own ethnic group	0.80	-0.01 (0.01)	-0.02 (0.04)	0.80	0.01 (0.01)	0.02 (0.04)
	1,545	3,292		1,311	2,895	
Trusts people in other ethnic groups	0.44	-0.03 (0.02)	-0.07 (0.05)	0.53	0.00 (0.02)	0.00 (0.06)
	1,545	3,286		1,310	2,871	
Believes elections are the best way to choose leaders	0.91	-0.00 (0.01)	-0.01 (0.03)	0.80	0.02 (0.02)	0.06 (0.05)
	1,537	3,278		1,211	2,658	
Believes boys are smarter than girls	0.35	-0.00 (0.01)	-0.00 (0.04)	0.51	0.02 (0.02)	0.07 (0.05)
	2,166	4,593		1,650	3,606	
<i>Panel B: Girls only</i>						
Trusts people in own ethnic group	0.79	-0.01 (0.02)	-0.02 (0.06)	0.80	0.00 (0.02)	0.01 (0.07)
	772	1,701		651	1,430	
Trusts people in other ethnic groups	0.43	-0.04 (0.03)	-0.10 (0.07)	0.49	0.02 (0.03)	0.05 (0.09)
	770	1,694		655	1,421	
Believes elections are the best way to choose leaders	0.89	0.01 (0.01)	0.01 (0.04)	0.78	0.01 (0.02)	0.04 (0.09)
	766	1,690		597	1,300	
Believes boys are smarter than girls	0.19	-0.02 (0.02)	-0.05 (0.04)	0.36	0.01 (0.02)	0.05 (0.08)
	1,074	2,329		821	1,787	
<i>Panel C: Boys only</i>						
Trusts people in own ethnic group	0.81	-0.01 (0.02)	-0.02 (0.06)	0.79	0.01 (0.02)	0.01 (0.06)
	773	1,591		660	1,465	
Trusts people in other ethnic groups	0.46	-0.02 (0.03)	-0.05 (0.07)	0.57	-0.02 (0.03)	-0.04 (0.07)
	775	1,592		655	1,450	
Believes elections are the best way to choose leaders	0.92	-0.01 (0.01)	-0.04 (0.04)	0.82	0.03 (0.02)	0.07 (0.05)
	771	1,588		614	1,358	
Believes boys are smarter than girls	0.51	0.02 (0.02)	0.04 (0.06)	0.66	0.03 (0.02)	0.08 (0.06)
	1,092	2,264		829	1,819	

Notes: Table reports the effects of the scholarship and attending Bridge on pupil social attitudes. Robust standard errors are reported in parentheses. All specifications include a linear control for the probability of treatment in the strata cell. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 20: Pupil and Caregiver (C) reported characteristics of schools

	Primary			Pre-primary		
	Bridge (1)	Public (2)	Private (3)	Bridge (4)	Public (5)	Private (6)
Fees (KES) (C)						
Total edu expenditure	13,631	5,886***	15,206***	11,126	5,320***	10,639
Fees (including meals)	8,787	2,111***	9,981***	7,426	2,746***	6,931*
Uniform	1,467	1,272***	1,769***	1,019	1,048	1,305***
Books	1,313	938***	1,549***	673	442***	734
Transportation	1,166	432***	1,116	1,333	547***	811***
Tutoring	492	597***	552	259	267	428***
Other	621	506***	498**	617	340***	459*
Observations	1,415	2,526	631	517	490	589
Facilities						
School has latrines	1.00	1.00	0.99***	0.99	0.98	0.99
Latrines are clean	0.87	0.83***	0.84*	0.83	0.80	0.80
Has library	0.38	0.57***	0.42*	0.42	0.46	0.39
Access to computer	0.28	0.62***	0.18***	0.35	0.33	0.20***
Pupil has used a computer	0.07	0.23***	0.07	0.11	0.14	0.12
Has a field for play	0.88	0.96***	0.88	0.92	0.92	0.88***
Field has hazards	0.36	0.32**	0.37	0.48	0.38***	0.34***
School has wall or fence	0.85	0.70***	0.73***	0.72	0.67**	0.60***
Observations	1,419	2,557	635	531	489	587
Corporal punishment						
Teachers carry sticks or canes	0.75	0.84***	0.80***	0.78	0.80	0.78
Teachers hit, slap, or pinch pupils	0.80	0.85***	0.84**	0.79	0.83*	0.77
Observations	1,419	2,559	635	527	489	587
Other characteristics						
Length of school day (hours)	9.0	8.3***	8.8***	7.4	6.0***	6.7***
Pupil attended some Saturday school	0.73	0.26***	0.39***	0.37	0.09***	0.18***
Missed class b/c unable to pay fees (C)	0.55	0.44***	0.53	0.54	0.50	0.57
Number days absent (past term)	2.37	2.10*	2.13	2.08	1.76	1.91
Study hours (usual day)	1.4	1.4	1.5			
Teacher is never absent from class	0.43	0.36***	0.49**	0.50	0.48	0.57**
Teacher uses local language	0.12	0.35***	0.16***	0.21	0.41***	0.27**
Classroom crowding index	-0.31	-0.07***	-0.21**	0.20	0.27	0.14
Instructional adaptation index	0.24	-0.06***	0.10***	0.16	-0.13***	0.02***
Teacher-classroom engagement index	0.09	-0.01***	0.07			
Observations	1,203	2,198	520	322	309	380

Notes: Table reports pupil and caregiver reported characteristics of school. Sample restricts to scholarship non-recipients to avoid risk of bias due to reciprocity effects. (C) indicates that data was given by caregivers. All other items were reported by pupils. The *instructional adaptation index* combines reports from pupils and parents on whether they feel that their teachers would respond if the pupil was progressing faster or slower than other students in the class. The actual number of observations in each group of characteristics may vary due to missingness. Length of school day is calculated based on pupil responses to school start and end times, and implausible cases outside one to eleven are set to missing. Values for pupil absences that exceed 30 days are set to 30 days, approximately the 99th percentile. *, **, and *** indicate that the difference between government or private school mean and Bridge mean is statistically significant at the 10%, 5% and 1% level respectively.

Table 21: Effects on pupil reported characteristics of school facilities

	Primary grades			Pre-primary		
	Y(0) (1)	Y(1) (2)	2SLS (3)	Y(0) (4)	Y(1) (5)	2SLS (6)
School has latrines	1.00 (0.01)	1.00 (0.00)	0.00 (0.01)	0.98 (0.01)	1.00 (0.01)	0.02* (0.01)
		4,541			3,699	
Latrines are clean	0.85 (0.03)	0.93 (0.01)	0.09*** (0.03)	0.82 (0.03)	0.86 (0.02)	0.04 (0.04)
		4,525			3,648	
School has a library	0.55 (0.03)	0.50 (0.03)	-0.06 (0.04)	0.41 (0.04)	0.55 (0.03)	0.14*** (0.05)
		4,499			3,412	
Access to computer	0.54 (0.03)	0.37 (0.02)	-0.17*** (0.04)	0.28 (0.03)	0.39 (0.03)	0.11*** (0.05)
		4,530			3,593	
Pupil has used a computer	0.17 (0.03)	0.09 (0.01)	-0.08*** (0.03)	0.20 (0.03)	0.11 (0.02)	-0.09*** (0.04)
		4,530			3,584	
School has a field for play	0.94 (0.01)	0.90 (0.02)	-0.05** (0.02)	0.88 (0.02)	0.90 (0.02)	0.02 (0.03)
		4,538			3,685	
Field has hazards	0.34 (0.04)	0.41 (0.03)	0.08* (0.04)	0.36 (0.04)	0.40 (0.04)	0.04 (0.05)
		4,169			3,321	
School has a wall or fence	0.60 (0.03)	0.66 (0.02)	0.06 (0.04)	0.57 (0.04)	0.72 (0.03)	0.15*** (0.04)
		4,532			3,642	

Notes: Table reports the effects of the scholarship and attending Bridge on pupil reported characteristics of school facilities. Y(0) indicates the expected outcome for compliers who do not attend Bridge using the procedure of [Abadie \(2002\)](#). Y(1) indicates the expected outcome for compliers who do attend Bridge. Sample size varies across outcomes because students who did not know the answer or refused to answer are treated as missing. *Field has hazards* is missing for pupils who report not having a field for play. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 22: Effects on instructional time

	Primary grades			Pre-primary		
	Non-recipient			Non-recipient		
	mean (1)	ITT (2)	2SLS (3)	mean (4)	ITT (5)	2SLS (6)
<i>School day length</i>						
Start time (hr of day)	7.33 2,094	0.01 (0.04) 4,379	0.03 (0.12)	7.85 1,211	-0.07 (0.07) 2,649	-0.19 (0.19)
End time (hr of day)	15.86 2,086	0.24*** (0.07) 4,376	0.66*** (0.20)	14.34 1,235	0.55*** (0.09) 2,690	1.52*** (0.25)
Length of school day (hrs)	8.62 1,817	0.21*** (0.06) 3,746	0.59*** (0.16)	6.71 1,087	0.61*** (0.10) 2,356	1.79*** (0.27)
<i>Saturday school</i>						
Pupil attended som Saturday school	0.37 2,150	0.19*** (0.01) 4,511	0.53*** (0.03)	0.21 1,594	0.10*** (0.01) 3,504	0.29*** (0.04)
Pupil attended Saturday school every week	0.24 2,152	0.14*** (0.01) 4,514	0.40*** (0.03)	0.10 1,651	0.08*** (0.01) 3,623	0.24*** (0.03)
<i>Teacher absence</i>						
Never absent in usual week	0.36 2,115	0.05*** (0.01) 4,427	0.14*** (0.04)	0.49 1,416	-0.01 (0.02) 3,107	-0.03 (0.05)
Never absent more than 4 hours	0.49 2,115	0.06*** (0.02) 4,427	0.16*** (0.04)	0.61 1,416	-0.01 (0.02) 3,107	-0.04 (0.05)
Never absent more than 8 hours	0.65 2,115	0.05*** (0.01) 4,427	0.14*** (0.04)	0.73 1,416	-0.02 (0.02) 3,107	-0.06 (0.05)
Never absent more than half week	0.87 2,115	0.02* (0.01) 4,427	0.05* (0.03)	0.91 1,416	-0.02 (0.01) 3,107	-0.05 (0.03)
Never leaves class alone	0.34 2,139	0.04*** (0.01) 4,492	0.12*** (0.04)	0.38 1,528	-0.01 (0.02) 3,331	-0.03 (0.05)

Notes: Table reports the effects of the scholarship and attending Bridge on the length of the school day, saturday school attendance, and teacher absences. All responses are pupil reported. Pupils were asked to report characteristics for their school or teachers in 2017. Note that teacher absence may not account for the use of substitute teachers. Length of school day is calculated based on pupil responses to school start and end times, and implausible cases outside one to eleven are set to missing. Sample size varies across outcomes because students who did not know the answer or refused to answer are treated as missing. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 23: Effects on measures of classroom crowding and disruption

	Primary grades			Pre-primary		
	Non-recipient			Non-recipient		
	mean (1)	ITT (2)	2SLS (3)	mean (4)	ITT (5)	2SLS (6)
Sometimes pupils do not have seats	0.18 2,158	-0.04*** (0.01) 4,539	-0.11*** (0.03)	0.24 1,674	-0.03** (0.01) 3,664	-0.09** (0.04)
Other pupils often bother the pupil	0.22 2,157	-0.04*** (0.01) 4,535	-0.10*** (0.03)	0.32 1,664	-0.01 (0.02) 3,651	-0.04 (0.04)
Other pupils sometimes bother the pupil	0.65 2,157	-0.05*** (0.01) 4,535	-0.15*** (0.04)	0.72 1,664	-0.02 (0.01) 3,651	-0.06 (0.04)
Other pupils sometimes fight	0.60 2,158	-0.06*** (0.01) 4,540	-0.18*** (0.04)	0.69 1,668	-0.01 (0.02) 3,662	-0.04 (0.04)
Outside noises make it hard to hear the tchr	0.45 2,158	-0.02 (0.01) 4,541	-0.05 (0.04)	0.52 1,656	-0.01 (0.02) 3,619	-0.02 (0.05)

Notes: Table reports the effects of the scholarship and attending Bridge on effects of measures of classroom crowding and disruption. All responses are pupil reported. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 24: Effects on access to textbooks

	Primary grades			Pre-primary		
	Non-recipient			Non-recipient		
	mean	ITT	2SLS	mean	ITT	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Pupil has at least one textbook	0.850	0.012 (0.011)	0.034 (0.029)	0.606	0.033** (0.017)	0.093** (0.047)
Pupil has at least two textbooks	0.794	0.022* (0.012)	0.061* (0.033)	0.522	0.047*** (0.017)	0.133*** (0.048)
# textbooks	3.530	0.230*** (0.055)	0.645*** (0.153)	2.031	0.169*** (0.068)	0.480*** (0.190)
<i>Subjects with textbook</i>						
English	0.757	0.032*** (0.013)	0.090*** (0.035)	0.508	0.031* (0.017)	0.089* (0.049)
Math	0.753	0.034*** (0.013)	0.096*** (0.035)	0.478	0.038** (0.017)	0.109** (0.049)
Kiswahili	0.742	0.022* (0.013)	0.062* (0.035)	0.468	0.061*** (0.017)	0.174*** (0.049)
Science	0.652	0.069*** (0.014)	0.194*** (0.038)	0.310	0.014 (0.016)	0.038 (0.045)
Social studies	0.626	0.072*** (0.014)	0.202*** (0.039)	0.267	0.025 (0.015)	0.070 (0.044)
Observations		4,534			3,425	

Notes: Table reports the effects of the scholarship and attending Bridge on access to textbooks. All responses are pupil reported. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 25: Effects on pedagogy, engagement, and feedback

	Primary grades			Pre-primary		
	Non-recipient			Non-recipient		
	mean (1)	ITT (2)	2SLS (3)	mean (4)	ITT (5)	2SLS (6)
Teacher uses languages besides English and Kiswahili	0.296	-0.068*** (0.013)	-0.190*** (0.035)	0.291	-0.039*** (0.014)	-0.112*** (0.041)
	2,162		4,547	1,635		3,593
Teachers explain things multiple ways if pupils do not understand	0.438	0.073*** (0.015)	0.205*** (0.043)	0.395	-0.042 (0.039)	-0.128 (0.116)
	2,128		4,484	392		806
Teachers ask questions to pupils to check understanding	0.503	0.055*** (0.016)	0.155*** (0.043)	0.449	-0.030 (0.040)	-0.090 (0.116)
	2,126		4,481	394		808
Teachers often asked questions of students	0.653	0.027 (0.019)	0.072 (0.048)			
	1,435		3,016			
Agrees that teacher only ask certain students questions	0.273	-0.026 (0.017)	-0.069 (0.043)			
	1,455		3,049			
Students often ask questions	0.472	0.046** (0.020)	0.119*** (0.050)			
	1,415		2,967			
Pupil liked how teacher treated them when they needed help	0.490	0.066*** (0.015)	0.186*** (0.042)	0.454	-0.024 (0.017)	-0.067 (0.048)
	2,157		4,536	1,622		3,557
Teachers are nice when pupils asks questions				0.903	0.004 (0.012)	0.012 (0.034)
				1,242		2,662
Teachers rewards students for good performance	0.913	0.007 (0.009)	0.020 (0.024)	0.812	0.003 (0.014)	0.009 (0.038)
	2,159		4,529	1,583		3,449
Teachers reward good performance with public recognition	0.511	-0.003 (0.015)	-0.010 (0.042)	0.394	0.031* (0.017)	0.087* (0.048)
	2,159		4,529	1,583		3,449
Pupil reports all teacher grading is helpful	0.307	0.014 (0.015)	0.039 (0.040)	0.315	0.009 (0.040)	0.030 (0.118)
	2,122		4,467	384		792
Teachers make time for play				0.738	0.015 (0.017)	0.043 (0.048)
				1,273		2,750

Notes: All responses are pupil reported. Survey items on the frequency with which teachers asked questions, whether teachers only asked certain students questions, and the frequency of student questions were only asked of pupils in Standard 4 and above. Survey items on teachers explaining things multiple ways, asking questions to check understanding, and grading were only asked of pupils in primary grades. Survey items on teachers being nice when pupils asks questions and whether teachers make time to play were only asked of students in pre-primary grades. Responses regarding rewards provided for good performance are classified indicating the use of public recognition if the pupil lists either *cheers* or *other public recognition* as rewards used by their teacher.

Table 26: Effects on test preparation, homework and study

	Primary grades			Pre-primary		
	Non-recipient			Non-recipient		
	mean (1)	ITT (2)	2SLS (3)	mean (4)	ITT (5)	2SLS (6)
Pupils affirms that they spend a lot of time practicing taking exams	0.68 2,121	0.02 (0.01) 4,465	0.06 (0.04)	0.62 382	0.06 (0.04) 784	0.17 (0.11)
Pupil affirms they spent a lot of class time preparing for a test	0.59 2,119	0.04*** (0.01) 4,449	0.12*** (0.04)	0.57 376	0.01 (0.04) 779	0.04 (0.12)
Pupil affirms they have taken a test like the one taken at endline	0.81 2,128	0.03*** (0.01) 4,475	0.09*** (0.03)	0.77 393	0.02 (0.04) 806	0.06 (0.10)
Days with homework in a usual week	3.4 2,165	0.6*** (0.1) 4,550	1.7*** (0.1)	3.6 1,698	0.3*** (0.1) 3,725	1.0*** (0.2)
Study hours (usual day)	1.4 1,455	0.1** (0.0) 3,051	0.2** (0.1)	1.3 51	0.3 (0.3) 87	0.9 (0.9)
Pupil has a quiet place to study at home	0.86 2,116	0.01 (0.01) 4,463	0.02 (0.03)	0.86 394	-0.01 (0.03) 810	-0.04 (0.09)
Pupil is assigned more work than other pupils in school	0.08 2,107	-0.00 (0.01) 4,442	-0.01 (0.02)	0.12 386	-0.01 (0.03) 794	-0.05 (0.08)
Pupil is assigned more work than pupils in other schools	0.19 1,929	0.05*** (0.01) 4,065	0.14*** (0.04)	0.21 353	0.00 (0.04) 733	0.00 (0.10)

Notes: Table reports the effects of the scholarship and attending Bridge on pupil perceptions of test preparation, homework, and study. All responses are pupil reported. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 27: Effects on parental engagement and school satisfaction

	Primary grades			Pre-primary		
	Non-recipient			Non-recipient		
	mean (1)	ITT (2)	2SLS (3)	mean (4)	ITT (5)	2SLS (6)
Member of household is on school committee	0.16	0.05*** (0.01)	0.14*** (0.03)	0.14	0.06*** (0.01)	0.18*** (0.04)
	2,177	4,559		1,672	3,666	
Caregiver knows pupil's class rank	0.72	0.04*** (0.01)	0.11*** (0.04)	0.73	0.07*** (0.01)	0.20*** (0.04)
	2,152	4,507		1,600	3,540	
Caregiver knows pupil's most recent test scores	0.78	0.01 (0.01)	0.02 (0.03)	0.78	0.05*** (0.01)	0.14*** (0.04)
	2,177	4,559		1,672	3,666	
Caregiver has met one of pupil's teacher	0.80	0.03*** (0.01)	0.10*** (0.03)	0.81	0.05*** (0.01)	0.15*** (0.04)
	2,173	4,548		1,660	3,648	
Caregiver has met pupil's head teacher'	0.80	0.03*** (0.01)	0.10*** (0.03)	0.81	0.05*** (0.01)	0.15*** (0.04)
	2,173	4,548		1,660	3,648	
Number meetings with pupil's teacher (past year)	3.3	0.6** (0.3)	1.7** (0.8)	4.0	1.4*** (0.4)	3.9*** (1.2)
	2,173	4,548		1,660	3,648	
Number meetings with head teacher (past year)	2.7	0.6** (0.3)	1.7** (0.8)	3.3	0.6 (0.4)	1.9 (1.3)
	2,171	4,537		1,665	3,648	
Parent is satisfied with pupil's school	0.71	0.07*** (0.01)	0.19*** (0.04)	0.71	0.06*** (0.01)	0.18*** (0.04)
	2,228	4,656		1,744	3,795	
Caregiver has complained to school	0.27	-0.02 (0.01)	-0.05 (0.04)	0.29	-0.03* (0.01)	-0.08* (0.04)
	2,090	4,390		1,644	3,611	
Caregiver has complained about fees	0.04	-0.00 (0.01)	-0.01 (0.02)	0.07	-0.03*** (0.01)	-0.09*** (0.02)
	2,090	4,390		1,644	3,611	

Notes: Table reports the effects of the scholarship and attending Bridge on caregiver reported engagement with school officials and knowledge of pupil performance. All responses are caregiver reported. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 28: Effects on corporal punishment

	Primary grades			Pre-primary		
	Non-recipient mean	ITT	2SLS	Non-recipient mean	ITT	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Teachers use harsh language or bad words	0.226 2,160	-0.018 (0.013) 4,543	-0.051 (0.035)	0.198 1,670	-0.002 (0.014)	-0.005 (0.039) 3,659
Teachers carry sticks or canes	0.812 2,163	-0.049*** (0.013) 4,548	-0.138*** (0.035)	0.785 1,682	-0.011 (0.014)	-0.033 (0.041) 3,689
Teachers hit, slap, or pinch students	0.835 2,164	-0.022* (0.012) 4,549	-0.061* (0.032)	0.803 1,681	0.000 (0.013)	0.000 (0.038) 3,687

Notes: Table reports the effects of the scholarship and attending Bridge on pupil reported experiences related to corporal punishment. All responses are pupil reported. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 29: Effects of scholarship and Bridge attendance on test conditions

	Primary			Pre-primary		
	Non-recipient	ITT	2SLS	Non-recipient	ITT	2SLS
	mean (1)	(2)	(3)	mean (4)	(5)	(6)
Pupil reports having a meal that day	0.82	0.00 (0.01)	0.01 (0.03)	0.85	0.00 (0.01)	0.01 (0.03)
	2,183	4,594		1,731	3,801	
Hours since last meal	6.09	-0.10 (0.19)	-0.28 (0.53)	5.46	-0.10 (0.19)	-0.28 (0.53)
	2,178	4,582		1,700	3,732	
Heavy rain/stormy weather	0.02	-0.00 (0.00)	-0.00 (0.01)	0.01	-0.00 (0.00)	-0.00 (0.01)
	2,294	4,818		1,813	3,970	
No desk or chair available	0.08	-0.00 (0.01)	-0.01 (0.02)	0.10	-0.00 (0.01)	-0.01 (0.02)
	2,294	4,818		1,813	3,970	
Loud noises during test	0.03	0.00 (0.01)	0.01 (0.01)	0.04	0.00 (0.01)	0.01 (0.01)
	2,294	4,818		1,813	3,970	
Siblings/other children disturbing	0.04	-0.00 (0.01)	-0.01 (0.01)	0.04	-0.00 (0.01)	-0.01 (0.01)
	2,294	4,818		1,813	3,970	
Low lighting	0.01	0.00 (0.00)	0.00 (0.01)	0.00	0.00 (0.00)	0.00 (0.01)
	2,294	4,818		1,813	3,970	
Desk shared by other pupil(s)	0.00	0.00** (0.00)	0.01** (0.00)	0.00	0.00** (0.00)	0.01** (0.00)
	2,294	4,818		1,813	3,970	
Other distractions	0.01	0.01* (0.00)	0.02* (0.01)	0.03	0.01* (0.00)	0.02* (0.01)
	2,294	4,818		1,813	3,970	
No distractions noted	0.80	0.00 (0.01)	0.00 (0.03)	0.76	0.00 (0.01)	0.00 (0.03)
	2,294	4,818		1,813	3,970	
Hour of day	13.28	-0.10 (0.08)	-0.29 (0.23)	13.28	-0.10 (0.08)	-0.29 (0.23)
	2,215	4,638		1,754	3,833	
F-stat from joint test		1.26			0.57	
P-value from joint test		0.24			0.86	

Notes: This table reports the effects of the scholarship and of Bridge attendance on FO reported testing conditions at the time of the endline survey. All specifications include controls for baseline characteristics, planned 2016 grade, and randomization strata fixed effects. The survey asked pupils to report the number of hours since their last meal. Where the reported value exceeds the hour of the day minus four, we impute the value to be the hour of day minus four. Pupils who report not having had a meal that day are coded as having had a meal at 5 pm the previous day. 88 respondents who refused to answer are coded as missing data on the number of hours since the last meal. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 30: Baseline balance table for the sibling instrument design

	P _{main}		PP _{main}	
	Non-recipient mean (1)	Coef. (2)	Non-recipient mean (4)	Coef. (5)
<i>Pupil characteristics</i>				
Pupil is female	0.52	-0.03	0.51	0.04
Age of pupil	10.52	-0.03	4.93	0.14
Mother is alive	0.94	0.02	0.98	-0.04***
Father is alive	0.82	-0.05*	0.86	-0.01
<i>Home characteristics</i>				
Floor of home is dirt	0.66	0.02	0.54	0.04
Walls of home are mud	0.54	0.02	0.46	0.02
Household has electricity	0.36	-0.02	0.37	0.01
Household has latrine	0.77	-0.03	0.71	-0.00
HH income (annual 000s KES)	42.33	-0.90	47.45	0.09
<i>Primary caregiver characteristics</i>				
Caregiver can read	0.12	-0.02	0.09	0.02
Caregiver can write	0.14	-0.01	0.12	0.04
<i>Caregiver occupation</i>				
Agriculture	0.15	0.00	0.11	-0.01
Casual laborer	0.44	0.01	0.42	-0.03
Small business owner	0.17	-0.04	0.18	-0.02
Observations	685	1,345	596	1,085
F-stat of joint test		1.44		0.99
P-value		0.13		0.47

Notes: This table describes the spillover design sample and compares the siblings of scholarship recipients and non-recipients in terms of baseline covariates. Sample is restricted to those pupils in the final control sample who were successfully interviewed at endline. The covariates reported in this table are from the application used to enroll the student in the scholarship program. Household income is winsorized at the 99th percentile. The coefficient estimates are from a test of the difference between siblings of scholarship recipients and non-recipients in the control group if more than 25 percent of pupil's siblings received the scholarship. The specification used controls for for planned 2016 grade and the number of pupils in the household. Standard errors are clustered at the household level. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 31: Effects of attending Bridge on non-recipients using sibling scholarship receipt instrument

	Primary				Pre-primary			
	Non-recipient				Non-recipient			
	Mean (1)	ITT (2)	2SLS (3)	N (4)	Mean (5)	ITT (6)	2SLS (7)	N (8)
<i>Panel A: First stage effect</i>								
Attended Bridge	0.18	0.12*** (0.03)		1,345	0.32	0.14*** (0.04)		1,085
<i>Panel B: Academic outcomes</i>								
Subject knowledge index	0.14	0.09 (0.12)	0.75 (1.04)	1,345	-5.24	0.32** (0.16)	2.22* (1.19)	1,085
<i>Panel C: Educational expenditures and missed class</i>								
Total household ed exp (past year, 000s)	31.27	1.17 (2.84)	9.96 (23.73)	1,345	31.30	0.97 (2.92)	6.78 (20.05)	1,085
Total ed exp on pupil (past year, 000s)	7.58	1.05* (0.61)	9.36* (4.98)	1,308	8.72	0.02 (0.64)	0.14 (4.52)	1,040
Fees (past year, 000s Kes)	3.73	0.55 (0.43)	4.95 (3.36)	1,260	5.41	0.39 (0.42)	2.74 (2.94)	1,002
Caregiver reports pupil missed class due to fees (past yr)	0.46	0.03 (0.04)	0.24 (0.36)	1,216	0.53	0.00 (0.05)	0.00 (0.32)	979
# days absent (past term, pupil report)	2.56	-0.07 (0.77)	-0.66 (6.85)	1,325	2.77	1.28 (1.04)	9.61 (7.98)	977
Tutoring (past year, 000s Kes)	0.52	-0.10 (0.10)	-0.85 (0.82)	1,274	0.35	-0.01 (0.11)	-0.09 (0.81)	1,019
Any tutoring (past yr)	0.29	-0.01 (0.04)	-0.09 (0.31)	1,294	0.18	-0.02 (0.04)	-0.12 (0.28)	1,011
# tutoring hrs per usual wk (past yr)	2.24	0.07 (0.40)	0.60 (3.32)	1,294	1.14	-0.27 (0.36)	-2.06 (2.83)	1,011

Notes: This table shows estimated effects of attending Bridge *at cost* among scholarship non-recipients using the first stage effect of scholarship assignment of other siblings. Outcome test scores are expressed in units of equivalent years of schooling (EYS). ITT effects are the effect of an indicator for over 25 percent of siblings receiving the scholarship. All specifications include controls for the number of pupils in the household. Sample is restricted to households with between 2 and 5 applicants to the scholarship program. The first stage relationship between this indicator and Bridge attendance of non-recipient siblings is given in Panel A. The ITT and 2SLS effects of Bridge attendance for non-recipient siblings on test score outcomes using this first stage effect are given in Panel B. Panel C shows analogous effects on educational expenditures and missed class. Standard errors are clustered at the level of the household/sibling group level. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 32: The effect of attending Bridge and fees paid on subject knowledge index
Primary

	OLS	2SLS			LIML	
Panel A: P_{main} (N=4,327)						
Attended Bridge school in 2017	0.54*** (0.04)	1.57*** (0.32) [18.59]	1.58*** (0.49) [17.75]	1.65*** (0.30) [15.15]	0.93*** (0.15) [8.51]	1.15*** (0.35)
School fees paid (000s KES)	0.02*** (0.00)	0.17*** (0.07) [3.23]	0.17 (0.10) [3.09]	0.18*** (0.06) [3.07]	0.04 (0.03) [2.31]	0.09 (0.08)
Kleibergen-Paap F-stat		2.27	1.85	2.27	1.69	
Panel B: P_{main} & P_{brig} (N=5,487)						
Attended Bridge school in 2017	0.71*** (0.04)	1.05*** (0.16) [40.41]	1.03*** (0.17) [111.99]	1.06*** (0.16) [30.04]	0.85*** (0.13) [10.38]	0.91*** (0.17)
School fees paid (000s KES)	0.03*** (0.00)	0.04** (0.02) [13.56]	0.04** (0.02) [40.68]	0.05*** (0.02) [11.14]	0.02 (0.02) [4.37]	0.03 (0.02)
Kleibergen-Paap F-stat		9.90	24.50	8.30	3.29	
Panel C: P_{main} & P_{priv} & P_{brig} (N=6,441)						
Attended Bridge school in 2017	0.68*** (0.03)	1.06*** (0.15) [38.56]	0.95*** (0.16) [114.28]	1.05*** (0.15) [29.53]	0.90*** (0.13) [9.78]	1.00*** (0.18)
School fees paid (000s KES)	0.03*** (0.00)	0.04*** (0.02) [14.48]	0.03* (0.02) [43.58]	0.05*** (0.02) [11.92]	0.04*** (0.01) [4.52]	0.05** (0.02)
Kleibergen-Paap F-stat		10.41	25.89	8.87	3.36	
<i>Instruments</i>						
Intention and demographic interactions		X		X	X	X
Academy-intention mprediction interactions			X	X	X	X
County dummies					X	X

Notes: The separate effects of attending Bridge and of paying fees are identified by multiple instruments formed by interacting scholarship assignment with a vector of baseline controls. Test score outcomes are expressed in units of equivalent years of schooling (EYS). School fees include tuition, meals, and uniform costs. All specifications include baseline characteristics, planned 2016 grade, and control linearly for the randomization strata probability of scholarship assignment. 2SLS specifications use instruments specified at the bottom of the table. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 33: The effect of attending Bridge and fees paid on subject knowledge index
Pre-primary

	OLS	2SLS			LIML	
Panel A: PP_{main} (N=3,575)						
Attended Bridge school in 2017	0.80*** (0.06)	1.92*** (0.41)	1.40** (0.67)	1.96*** (0.40)	1.34*** (0.24)	1.53*** (0.43)
School fees paid (000s KES)	0.02*** (0.01)	0.07 (0.05)	0.00 (0.09)	0.07 (0.05)	0.01 (0.03)	0.02 (0.05)
Kleibergen-Paap F-stat		3.03	2.73	2.64	2.22	
Panel B: PP_{main} & PP_{brig} (N=3,950)						
Attended Bridge school in 2017	0.87*** (0.05)	2.01*** (0.36)	1.90*** (0.43)	2.00*** (0.33)	1.52*** (0.24)	1.96*** (0.46)
School fees paid (000s KES)	0.02*** (0.00)	0.08** (0.04)	0.07 (0.05)	0.08** (0.04)	0.03 (0.02)	0.08 (0.05)
Kleibergen-Paap F-stat		3.89	7.32	3.85	2.10	
<i>Instruments</i>						
Intention and demographic interactions		X		X	X	X
Academy-intention mprediction interactions			X	X	X	X
County dummies					X	X

Notes: The separate effects of attending Bridge and of paying fees are identified by multiple instruments formed by interacting scholarship assignment with a vector of baseline controls. Test score outcomes are expressed in units of equivalent years of schooling (EYS). School fees include tuition, meals, and uniform costs. All specifications include baseline characteristics, planned 2016 grade, and control linearly for the randomization strata probability of scholarship assignment. 2SLS specifications use instruments specified at the bottom of the table. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 34: The effect of attending Bridge and fees paid on subject knowledge index
Enrollment variable: years of Bridge

	OLS	2SLS			LIML	
Panel A: P_{main} (N=4,327)						
Years enrolled at Bridge	0.27*** (0.02)	0.70*** (0.14) [26.17]	0.72*** (0.22) [23.50]	0.74*** (0.13) [21.06]	0.47*** (0.07) [11.08]	0.54*** (0.13)
School fees paid (000s KES)	0.02*** (0.00)	0.14** (0.06) [3.80]	0.15 (0.10) [3.44]	0.16*** (0.05) [3.53]	0.04 (0.03) [2.43]	0.07 (0.06)
Kleibergen-Paap F-stat		2.72	2.08	2.65	1.92	
Panel B: P_{main} & P_{brig} (N=5,194)						
Years enrolled at Bridge	0.26*** (0.02)	0.48*** (0.07) [59.30]	0.47*** (0.08) [159.44]	0.48*** (0.07) [43.68]	0.42*** (0.06) [13.90]	0.43*** (0.08)
School fees paid (000s KES)	0.02*** (0.00)	0.03 (0.02) [15.92]	0.03 (0.02) [46.87]	0.03* (0.02) [12.86]	0.02 (0.01) [4.73]	0.02 (0.02)
Kleibergen-Paap F-stat		12.60	30.45	10.37	3.88	
Panel C: P_{main} & P_{priv} & P_{brig} (N=6,112)						
Years enrolled at Bridge	0.26*** (0.02)	0.48*** (0.07) [55.37]	0.44*** (0.07) [159.32]	0.48*** (0.07) [42.06]	0.44*** (0.06) [13.08]	0.46*** (0.07)
School fees paid (000s KES)	0.02*** (0.00)	0.03* (0.02) [17.17]	0.02 (0.02) [50.80]	0.03** (0.01) [13.92]	0.03** (0.01) [5.04]	0.03* (0.02)
Kleibergen-Paap F-stat		13.40	32.42	11.23	4.03	
<i>Instruments</i>						
Intention and demographic interactions		X		X	X	X
Academy-intention mprediction interactions			X	X	X	X
County dummies					X	X

Notes: The separate effects of attending Bridge and of paying fees are identified by multiple instruments formed by interacting scholarship assignment with a vector of baseline controls. Test score outcomes are expressed in units of equivalent years of schooling (EYS). School fees include tuition, meals, and uniform costs. All specifications include baseline controls, a linear control for the probability of treatment with the strata. 2SLS specifications use instruments specified at the bottom of the table. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 35: The effect of attending Bridge and fees paid on subject knowledge index
Enrollment variable: years of Bridge

	OLS	2SLS			LIML	
Panel A: PP_{main} (N=3,575)						
Years enrolled at Bridge	0.43*** (0.03)	0.92*** (0.19)	0.66** (0.31)	0.95*** (0.19)	0.65*** (0.12)	0.70*** (0.21)
		[12.14]	[14.05]	[9.67]	[23.56]	
School fees paid (000s KES)	0.02*** (0.01)	0.06 (0.05)	-0.01 (0.08)	0.07 (0.05)	0.01 (0.03)	0.01 (0.05)
		[4.73]	[5.93]	[4.03]	[8.51]	
Kleibergen-Paap F-stat		3.38	3.29	2.95	5.48	
Panel B: PP_{main} & PP_{brig} (N=3,726)						
Years enrolled at Bridge	0.41*** (0.03)	0.96*** (0.16)	0.87*** (0.18)	0.95*** (0.15)	0.73*** (0.11)	0.88*** (0.20)
		[16.37]	[41.71]	[14.40]	[24.57]	
School fees paid (000s KES)	0.01*** (0.00)	0.07* (0.04)	0.05 (0.04)	0.07** (0.03)	0.03 (0.02)	0.06 (0.05)
		[6.68]	[17.74]	[6.48]	[9.24]	
Kleibergen-Paap F-stat		4.65	9.81	4.65	5.65	
<i>Instruments</i>						
Intention and demographic interactions		X		X	X	X
Academy-intention mprediction interactions			X	X	X	X
County dummies					X	X

Notes: The separate effects of attending Bridge and of paying fees are identified by multiple instruments formed by interacting scholarship assignment with a vector of baseline controls. Test score outcomes are expressed in units of equivalent years of schooling (EYS). School fees include tuition, meals, and uniform costs. All specifications include baseline controls, a linear control for the probability of treatment with the strata. 2SLS specifications use instruments specified at the bottom of the table. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 36: Effect on grade progression with multiple endogenous variables

	Primary grades			Pre-primary grades		
	OLS	2SLS		OLS	2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)
Separate estimation of Bridge enrollment and fee effects						
Enrolled at Bridge	0.102*** (0.016)	0.238*** (0.048)	0.252*** (0.046)	0.108*** (0.021)	0.133 (0.096)	0.226*** (0.083)
		[113.353]	[38.559]		[26.832]	[12.335]
School fees paid (000s KES)	0.004*** (0.001)	0.007 (0.005)	0.009** (0.005)	0.002 (0.002)	-0.004 (0.011)	0.005 (0.009)
		[43.140]	[14.467]		[15.227]	[6.374]
Kleibergen-Paap F-stat		25.600	10.407		7.563	4.035
Observations		6,113	6,442		3,722	3,946
<i>Instruments</i>						
Academy-intention prediction interactions		X			X	
Intention and demographic interactions			X			X

Notes: Table reports results from specifications that separately estimate the effects of enrolling at Bridge and of paying fees. School fees include tuition, meals, and uniform costs. Column (1) reports a non-experimental OLS specification restricting the sample to the control group. All 2SLS specifications are identified using multiple instruments formed by interacting scholarship assignment with a vector of baseline controls. *Academy-intention predictions* are four instruments representing the empirical likelihood that a pupil in a particular academy-intention cell will enroll in a public school, Bridge school, non-Bridge private school, or remain unenrolled in the absence of the scholarship. Each prediction is the empirical Bayes shrunk leave-i-out share of pupils choosing each enrollment type among scholarship non-recipients. The *intention and demographic interactions* include the following ten indicators: being enrolled in any school at time of application; being enrolled in NGO/private school; any prior Bridge enrollment; having plans to enroll in a Bridge school/public school/non-Bridge private school in the absence of the scholarship; primary caregiver's is a agricultural/casual laborer; whether the household applied through IPA; and Inverse Hyperbolic Sine (IHS) transformation of reported household monthly income. All specifications include baseline characteristics, planned 2016 grade, and control linearly for the randomization strata probability of scholarship assignment. 2SLS specifications use instruments specified at the bottom of the table. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 37: Separate effects of private and public school attendance on subject knowledge index

	OLS	2SLS			LIML	JIVE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: PP_{main} (N=3,811)							
Years attending public school	-0.73*** (0.07)	-2.10*** (0.68)	-1.07 (1.45)	-2.04*** (0.64)	-1.91*** (0.38)	-3.27** (1.46)	-0.96* (0.51)
		[3.92]	[2.80]	[3.31]	[2.77]		
Years attending private school	-0.55*** (0.06)	-1.10** (0.56)	-1.80* (1.08)	-1.13** (0.50)	-0.90*** (0.33)	0.03 (1.16)	-2.12*** (0.42)
		[6.44]	[5.28]	[5.92]	[4.33]		
Kleibergen-Paap F-stat		1.71	1.05	1.62	1.50		
Panel B: P_{main} (N=4,621)							
Years attending public school	-0.49*** (0.04)	-1.09*** (0.20)	-0.68*** (0.18)	-0.83*** (0.15)	-0.69*** (0.13)	-0.50 (0.32)	-0.92*** (0.15)
		[34.45]	[102.78]	[40.87]	[12.07]		
Years attending private school	-0.25*** (0.09)	1.15 (1.48)	-2.14* (1.24)	-0.73 (0.87)	-1.25* (0.66)	-3.31 (2.84)	0.17 (0.73)
		[2.00]	[5.82]	[3.00]	[2.21]		
Kleibergen-Paap F-stat		1.47	3.53	2.33	1.23		
Panel C: P_{main} & P_{priv} (N=5,612)							
Years attending public school	-0.57*** (0.04)	-0.96*** (0.13)	-0.78*** (0.14)	-0.88*** (0.13)	-0.76*** (0.12)	-0.78*** (0.15)	-0.63 (1.61)
		[57.62]	[156.28]	[44.46]	[12.78]		
Years attending private school	-0.11* (0.07)	-0.01 (0.49)	-0.89* (0.52)	-0.20 (0.45)	-0.50 (0.40)	-0.58 (0.72)	-1.95 (10.88)
		[5.80]	[16.67]	[5.18]	[2.41]		
Kleibergen-Paap F-stat		4.82	11.39	4.51	1.99		
Panel D: P_{main} & P_{priv} & P_{brig} (N=6,816)							
Years attending public school	-0.72*** (0.04)	-0.96*** (0.13)	-0.78*** (0.14)	-0.87*** (0.13)	-0.75*** (0.12)	-0.77*** (0.16)	0.52 (7.49)
		[52.48]	[158.14]	[42.09]	[12.50]		
Years attending private school	-0.28*** (0.06)	0.18 (0.48)	-0.53 (0.50)	-0.09 (0.44)	-0.32 (0.38)	-0.28 (0.68)	-7.88 (42.54)
		[6.18]	[18.27]	[5.31]	[2.77]		
Kleibergen-Paap F-stat		5.06	12.24	4.55	2.10		
<i>Instruments</i>							
10 covariates		X		X	X	X	X
4 LOO predictions			X	X	X	X	X
44 county dummies					X	X	X

Notes: This table presents results from tests that separately identify the effect of public and private school relative the counterfactual of Bridge attendance. Test score outcomes are expressed in units of equivalent years of schooling (EYS). All specifications include baseline characteristics, planned 2016 grade, and control linearly for the randomization strata probability of scholarship assignment. 2SLS specifications use instruments specified at the bottom of the table. Angrist-Pischke F-stats are reported in brackets. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 38: Separate effects of private and public school attendance on subject knowledge index
 Endogenous variables: Years of attendance

	OLS	2SLS			LIML	JIVE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: PP_{main} (N=3,811)							
Years Attending government school	-0.40*** (0.04)	-1.17*** (0.35) [4.53]	-0.63 (0.68) [4.04]	-1.16*** (0.34) [3.86]	-0.97*** (0.21) [4.75]	-1.51** (0.65)	-0.39 (0.36)
Years Attending private school	-0.25*** (0.03)	-0.53** (0.25) [8.48]	-0.84* (0.45) [9.69]	-0.52** (0.23) [8.01]	-0.47*** (0.16) [18.33]	-0.16 (0.46)	-1.08*** (0.27)
Kleibergen-Paap F-stat		2.16	1.72	2.11	2.56		
Panel B: P_{main} (N=4,621)							
Years Attending government school	-0.26*** (0.02)	-0.52*** (0.11) [29.01]	-0.34*** (0.08) [129.23]	-0.39*** (0.08) [47.06]	-0.34*** (0.06) [14.82]	-0.30*** (0.09)	-0.51*** (0.10)
Years Attending private school	-0.12*** (0.05)	0.42 (0.92) [1.70]	-0.98* (0.55) [8.82]	-0.59 (0.46) [3.46]	-0.74** (0.33) [2.44]	-1.29* (0.73)	0.55 (0.63)
Kleibergen-Paap F-stat		1.22	5.18	2.59	1.51		
Panel C: P_{main} & P_{priv} (N=5,612)							
Years Attending government school	-0.31*** (0.02)	-0.46*** (0.06) [69.55]	-0.37*** (0.07) [196.84]	-0.42*** (0.06) [55.03]	-0.38*** (0.06) [16.33]	-0.37*** (0.07)	-0.29 (0.31)
Years Attending private school	-0.05 (0.04)	-0.08 (0.27) [5.40]	-0.50* (0.26) [19.13]	-0.23 (0.23) [5.45]	-0.37* (0.20) [2.23]	-0.47 (0.31)	-1.32 (2.47)
Kleibergen-Paap F-stat		4.56	13.44	4.80	2.02		
Panel D: P_{main} & P_{priv} & P_{brig} (N=6,816)							
Years Attending government school	-0.40*** (0.02)	-0.47*** (0.06) [62.09]	-0.38*** (0.07) [198.20]	-0.42*** (0.06) [51.67]	-0.37*** (0.06) [15.33]	-0.37*** (0.07)	-0.31 (0.32)
Years Attending private school	-0.16*** (0.03)	-0.01 (0.26) [6.02]	-0.36 (0.25) [20.25]	-0.19 (0.23) [5.64]	-0.31 (0.20) [2.35]	-0.39 (0.31)	-0.90 (2.14)
Kleibergen-Paap F-stat		5.07	14.12	4.96	2.09		
<i>Instruments</i>							
10 covariates		X		X	X	X	X
4 LOO predictions			X	X	X	X	X
44 county dummies					X	X	X

Notes: This table presents results from tests that separately identify the effect of public and private school relative the counterfactual of Bridge attendance. The endogenous variables represent the *number of years of attendance* at each school type. Test score outcomes are expressed in units of equivalent years of schooling (EYS). All specifications include baseline characteristics, planned 2016 grade, and control linearly for the randomization strata probability of scholarship assignment. 2SLS specifications use instruments specified at the bottom of the table. Angrist-Pischke F-stats are reported in brackets. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 39: Implied weights on sub-LATEs

	Public (1)	Private (2)	Unenrolled (3)	Bridge <i>at cost</i> (4)
Primary school	0.58	0.06	<0.00	0.34
Pre-primary	0.24	0.29	0.02	0.44

Notes: Table shows the implied weights on each sub-LATE given the first stage effects on attendance. Weights are given by the negative ratio of the first stage effect on the counterfactual and the first stage effect on attending Bridge on scholarship. The results in this table are from a sample that restricts to the P_{main} and PP_{main} samples.

Table 40: Relationship between endline scores and Bridge internal scores

	Primary			Pre-primary		
	English (1)	Kiswahili (2)	Math (3)	English (4)	Kiswahili (5)	Math (6)
Coefficient	0.49***	0.51***	0.42***	0.41***	0.49***	0.38***
SE	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)
R^2	0.24	0.23	0.22	0.15	0.20	0.13
Observations	925	925	907	822	418	808

Notes: Table reports the relationship between Bridge's internal test score records and the endline outcomes. Each row represents the results of regressing the endline subject test score on Bridge's test score. The endline scores are not equated to a single scale in order to be symmetric with the Bridge scores, which are standardized within grade. We use the pupil's 2017 term 3 midterm exam for the internal Bridge test score. The sample is constructed by matching scholarship recipients in the PP_{main} and P_{main} samples to Bridge test score data using an crosswalk provided by Bridge. The file was maintained for auditing purposes for the donor that provided the scholarships.***, **, and * indicate significance at 1%, 5%, and 10%.

Table 41: Effect of a year of exposure of Bridge on internal Bridge test scores - Duration exposure test

	Pooled subjects		English		Kiswahili		Math	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A: 200 day exposure window</i>								
Years exposure	0.589*** (0.058)	0.588*** (0.064)	0.556*** (0.076)	0.546*** (0.106)	0.421*** (0.086)	0.406*** (0.121)	0.742*** (0.091)	0.759*** (0.126)
Years exposure × Scholarship	-0.013 (0.091)	-0.025 (0.100)	0.092 (0.117)	0.100 (0.163)	-0.108 (0.132)	-0.116 (0.184)	-0.060 (0.135)	-0.097 (0.188)
Scholarship	0.038 (0.045)		-0.011 (0.050)		0.059 (0.060)		0.075 (0.056)	
# test scores		8,800		3,352		2,238		3,210
# pupils		1,841		1,676		1,120		1,605
<i>Panel B: 365 day exposure window</i>								
Years exposure	0.338*** (0.034)	0.338*** (0.037)	0.300*** (0.044)	0.298*** (0.054)	0.237*** (0.048)	0.230*** (0.059)	0.448*** (0.050)	0.455*** (0.061)
Years exposure × Scholarship	0.031 (0.051)	0.026 (0.055)	0.186*** (0.066)	0.189*** (0.080)	-0.074 (0.073)	-0.076 (0.089)	-0.061 (0.072)	-0.077 (0.088)
Scholarship	0.029 (0.042)		-0.029 (0.047)		0.055 (0.057)		0.073 (0.051)	
# test scores		13,201		5,028		3,358		4,815
# pupils		1,841		1,676		1,120		1,605
<i>Panel C: 600 day exposure window</i>								
Years exposure	0.238*** (0.021)	0.334*** (0.027)	0.310*** (0.024)	0.361*** (0.036)	0.142*** (0.029)	0.176*** (0.041)	0.233*** (0.028)	0.410*** (0.042)
Years exposure × Scholarship	0.037 (0.029)	0.036 (0.030)	0.031 (0.034)	0.037 (0.037)	0.036 (0.040)	0.035 (0.044)	0.043 (0.038)	0.037 (0.042)
Scholarship	0.024 (0.041)		0.014 (0.044)		0.026 (0.055)		0.035 (0.048)	
# test scores		22,008		8,379		5,605		8,024
# pupils		1,841		1,676		1,124		1,605
Pupil FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: This table shows the results of the effect of duration exposed to Bridge schools on Bridge's internal test score outcomes using a balanced panel of pupil mid-term test scores. The test score data were provided by Bridge. The table restricts to pupils who entered Bridge between January 1 and February 1, 2016 and recorded at least one test in the first 100 days of enrollment. Years of exposure is the number of days since enrolling divided by 365. The interaction term indicates the difference in the effect for pupils who arrived on scholarship. Columns 1 and 2 show results that stack each of the three subject scores. All specifications include grade fixed effects. The pooled specifications include interactions between grade and subject fixed effects. Standard errors are clustered at the pupil level for all specifications. ***, **, and * indicate significance at 1%, 5%, and 10%.

Table 42: Bridge academy staff characteristics

	Employed as of August 2016				Employed as of February 2019			
	Teachers		Academy managers	Substitute teachers	Teachers		Academy managers	Substitute teachers
	Primary (1)	Pre-primary (2)			Primary (5)	Pre-primary (6)		
Female	0.54 [0.50]	0.78 [0.41]	0.32 [0.47]	0.61 [0.49]	0.55 [0.50]	0.75 [0.43]	0.37 [0.48]	0.51 [0.50]
Age	27.39 [5.28]	29.02 [5.93]	29.72 [6.21]	27.81 [5.66]	27.79 [4.46]	28.86 [5.54]	30.13 [5.19]	28.12 [4.46]
Teaching certificate	0.23 [0.42]	0.13 [0.34]	0.12 [0.32]	0.18 [0.39]	0.57 [0.50]	0.63 [0.48]	0.24 [0.43]	0.61 [0.49]
Teaching Diploma	0.00 [0.00]	0.00 [0.03]	0.00 [0.05]	0.00 [0.02]	0.00 [0.03]	0.01 [0.09]	0.00 [0.05]	0.00 [0.00]
Bachelors degree	0.00 [0.00]	0.00 [0.00]	0.01 [0.07]	0.00 [0.00]	0.00 [0.02]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]
Masters degree	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]	0.00 [0.00]				
Years working at Bridge	1.78 [1.03]	2.15 [1.03]	1.89 [0.97]	1.26 [0.68]	1.61 [0.87]	1.30 [0.87]	1.80 [0.84]	1.86 [0.95]
First year teaching at Bridge	0.31 [0.46]	0.19 [0.39]	0.24 [0.43]	0.47 [0.50]	0.30 [0.46]	0.43 [0.50]	0.25 [0.43]	0.28 [0.45]
Monthly salary (KES)	10,143 [1,542]	10,102 [1,535]	11,494 [913]	10,290 [1,567]				
Observations	3,037	1,162	398	1,840	2,479	550	343	344

Notes: This table reports mean characteristics of teachers and academy managers according to Bridge administrative data. The results for teachers restrict to permanent teachers. We drop permanent teachers who are missing a class assignment (4 in 2015 and 8 in 2019). Standard deviations are reported in brackets.

Table 43: Characteristics of Bridge pupils and classrooms

	Primary (1)	Pre-primary (2)
Female	0.48	0.49
Years enrolled at Bridge	2.07	1.11
Pupil-teacher ratio (mean across pupils)	29.25	13.15
Pupil-teacher ratio (mean across classrooms)	21.28	8.07
Observations	69,496	8,896

Notes: This table reports mean characteristics of pupils and classrooms as Bridge academies. The results are based on historical enrollment records maintained by Bridge and provided to the research team. The data include initial enrollment date, withdrawal date, academy, grade, and stream for all pupils enrolled in Bridge between 2013 and 2019. The results in this table report characteristics of Bridge pupils that were enrolled in Bridge as of January 2017.