Incentivizing Textbooks for Self-Study: Experimental Evidence on Student Learning from the Democratic Republic of Congo*

Jean-Benoit Falisse [1] Marieke Huysentruyt [2,3] Anders Olofsgård [4, 5]

Date of this version: May 2019

Abstract

We use a randomized field experiment to test whether a new self-study routine, designed to encourage the use of textbooks at home, can improve student achievement. In treatment schools, students and teachers were incentivized to adopt the routine through, respectively, a public display of stars (one for each time they took home books) and financial incentives (to compensate for potential loss or damage of textbooks). French language test scores improved in the treatment schools by 0.319σ relative to the control group, but no impact on math test scores was found. The intervention also raised the average likelihood of a student taking the high-stake end of the year national exam by 10 percentage points, without a negative impact on average exam results. The routine increased students' job and secondary school aspirations and their perceptions of the usefulness of textbooks, likely pathways for the main results. The low-cost routine relied on a more efficient usage of existing basic educational material, making it feasible also in a very resource constrained and fragile setting. Our findings highlight the critical role of self-learning to promote student achievement and suggest that a simple 'textbook at home' routine may compensate for lower quality teaching in class.

JEL codes: C93, I21, J24, O15

* We are deeply grateful for the support from Cordaid to facilitate this research. We are particularly indebted to Adolphe Malanga, Emmanuel Muzi, Liliane Rwezangabo Milimbo, Kamala Kaghoma and Patrice Kiziba Yafali for excellent project management, supervision of surveys and data gathering, and to Alinda Bosch, Mamadou Sylla, Christian Bahizi and Manuela Jansen for continuous support and facilitation. We also thank Bluesquare (and especially Julie Vanhamme and Antoine Legrand) for their assistance with the RBF data and the surveys. We are also grateful for comments and suggestion from Abhijeet Singh, Pamela Campa, and Dagmara Celik Katreniak. Excellent research assistance was provided by Edvard Von Sydow. Finally, this research was made possible by the generous financial support from the World Bank, through the Results in Education for All Children (REACH) program.

[1] University of Edinburgh, Centre of African Studies (jb.falisse@ed.ac.uk).
[2] HEC Paris, Strategy and Business Policy (huysentruyt@hec.fr).
[3] Stockholm Institute of Transition Economics (SITE).
[4] Stockholm Institute of Transition Economics (SITE).
[5] Stockholm School of Economics (anders.olofsgard@hhs.se).

1 Introduction

Over the past three decades, numerous nationwide programs and billions of donor funds have focused on making textbooks widely available at primary schools, even in the world's poorest and most fragile settings like the Democratic Republic of Congo (World Bank, 2018; Fredriksen et al., 2015; Read 2015). Yet, the evidence on the impact of textbooks on student achievement is at best equivocal and mostly disappointing (Glewwe and Muralidharan, 2016; Glewwe et al., 2009; Sabarwal et al., 2014). One striking observation is that these textbook distribution programs rarely provide incentives for teachers and students to actually use the textbooks, and that school textbooks are seldom leveraged as a tool to encourage student self-study at home. This begs the important policy question of whether incentives are needed for the sunk investment in textbooks to really pay-off. Might the introduction of incentives that encourage greater utilization of textbooks offer a valuable, cost-effective approach towards improving students' learning outcomes and aspirations for the future?

This paper presents experimental evidence on the impact on student achievement and aspirations from a new incentivized routine encouraging students to bring textbooks home for self-study. To conduct the field experiment, we collaborated with Cordaid (Caritas Netherlands), one of the largest development aid organizations in the Netherlands, which has been active in the education sector in the South Kivu province of the Democratic Republic of Congo since 2008. In the aftermath of Africa's 'First World War,' which led to the loss of some five million lives between 1994 and 2003, the education sector as much else was in desperate need of reconstruction. The DRC government and the donor community, led by the World Bank, has invested substantial amounts in rebuilding some of that capacity, including a large textbook distribution program that has allocated around 40 million textbooks country-wide since 2008. Yet, it appears that the textbooks are commonly left unused or used irregularly and only in the classroom.

The experiment that we designed encouraged 5th and 6th grade students from 45 randomly selected schools to regularly take home textbooks and use them to study for weekly quizzes. We used a mix of financial and non-financial incentives mostly for students: (i) a public display of stars assigned

to each student that brought math and French textbooks home and back in good condition and participated in the end-of-week quiz, (ii) an in kind gift of pens and pencils for all students in classes where at least 75 % of the students with regularity participated in the routine, and (iii) a flat small compensation to schools for participating in the intervention and to compensate for a potential increase in the number of lost and damaged books. This new routine leverages textbooks as one of the main widely available pedagogical resources, fits well with the existing habits of self-study and in-class quizzes, and deliberately targets students instead of teachers whose motivation and qualifications in our context are notoriously low. The main goal of the intervention is twofold: to increase students' learning at home and thus improve their achievement, and to affect perceptions about the value of home study and education more broadly and thus shift students' beliefs and aspirations.

To measure student achievement, we rely on self-conducted test results in the French language and math, as well as the scores in the high-stake national exam (TENAFEP) that determine eligibility to secondary education. We emphasize three main sets of results. First, we find that the students in the treatment schools scored 0.28 to 0.32σ higher in French language tests relative to students in the control group schools. We evidence better performance both on the sentence completion tasks and the retrieving words tasks. However, the intervention had no significant impact on student test scores in math in aggregate, nor on any of the three main competency domains in math tested. These are intention-to-treat (ITT) estimates, reflecting an average compliance rate of over 80 percent. The local average treatment effect (LATE) on French and math test results were as expected somewhat larger, as high as 0.36σ for French but still statistically insignificant for math.

Second, we find that the intervention significantly raised, by 10 percent, the number of students taking the high stakes national exam and passing the test. That the intervention led more students to pass the exam is encouraging, as it is necessary for students to proceed to the next level of education. This increase in the number of students taking and passing the test was also achieved without a negative impact on average test scores, even though the additional students were likely drawn from the lower end of the skills distribution.

Third, we also sought to shed light on the impact of the 'textbook for home-study' routine on students' habits, attitudes and aspirations related to learning more broadly. We find that the intervention led students to express a more positive attitude towards textbooks, a greater wish to continue to secondary school and more ambitious job aspirations. These effects help explain why the intervention succeeded not only in raising student achievement but also in encouraging more students to take the TENAFEP, a prerequisite for continued studies and most non-manual jobs. Further, by affecting students' mindsets about studying and future careers, the consequences of the intervention may well be long-lasting.

This intervention differs from most others evaluated in the development economics literature in at least two ways. It focuses on more efficient usage of existing inputs rather than provision of new resources, and the incentives are focused on the students rather than the teachers and in a way that doesn't give raise to incentives to cheat or teaching to the test. Behrman et al. (2015), evaluating monetary performance incentives linked to test results across 88 Mexican high schools, found that treatment arms with incentives on students only and incentives on both students, teachers and administrators positively impacted test scores.[6] They also found, though, as confirmed in Martinelli et al. (2018), that the incentives gave rise to substantial cheating from the students. Similarly, it has been shown that teacher incentives based on student test results can also give raise to cheating (Jacob and Levitt, 2003).

The simple incentive design developed here can be beneficial everywhere, but perhaps in particular so in severely resource constrained and fragile settings such as DRC. Indeed, even though the immediate economic and social repercussions of the 'learning crisis' are particularly evident in fragile settings (World Bank, 2018), we know surprisingly little about how student achievement can be effectively raised under multiple and severe (resource) constraints, including particularly weak teacher capacity. Rigorous randomized experiments are rarely seen in fragile or conflict-affected states (Burde et al. 2017): of the 118 evaluations of education interventions reviewed by Glewwe and Muralidharan

[6] Kremer et al (2009) also look at student incentives and find a positive impact on test scores from a merit scholarship program for sixth-grade girls in Kenya.

(2016), only two were conducted in any of the 15 countries deemed most fragile according to the Fragile States Index produced by the think tank Fund for Peace.[7] Further, our study gives evidence of a 'low-hanging fruit'-type intervention: the intervention does not require any new inputs, instead it leverages an existing widespread yet underutilized educational resource to encourage self-study. Our findings suggest that such an intervention represents a valuable complement (not substitute) to the more difficult, slower-paced investments in teacher and school leadership capacity.

Implemented at small scale, the intervention compares favorably in terms of cost effectiveness, relative to interventions that increase teaching resources or focus on improved pedagogy (Kramer et al., 2013). Based on per student expenditures and the most conservatively estimated ITT effect of 0.28σ, we estimate that US$100 yields a 1.6σ improvement in test scores, or alternatively that US$63 is necessary to achieve a 1σ improvement. If scaled up to schools where a results-based financing (RBF) scheme exists[8], then the intervention's cost-effectiveness is likely to remain the same or even improve. Drawing on prior research, a bundle of financial and non-financial incentives was introduced in the pilot-test to maximize manipulation strength and thus ensure that our intervention is sufficiently high powered. However, a better understanding of the relative importance of these incentives may well suggest ways to further lower costs. In fact, given that few books actually went missing or were severely damaged, the fixed school-level transfer to compensate for such loss and damage may not need to be so high. On the other hand, if scaled up to schools without RBF, then actual costs of incentives and monitoring may well be higher.

This paper relates to the literature on student achievement in developing countries in general, and on the role of incentives, inputs and combinations thereof to student self-study and learning

[7] Burde and Linden (2013) evaluated the impact of building schools in rural Afghanistan on school enrollment and student achievement. They found large effects on average student test scores (specifically, an increase of 0.4 σ for boys and 0.6 σ for girls). However, though the authors do not provide an analysis of the project's cost effectiveness, the costs of this intervention are likely also high. Orkin (2013) assessed the impact of an increase in instructional time on student learning outcomes in Ethiopia and found that this increase had very small effects on student achievement. Again, no data on the implementation costs of the intervention were provided.

[8] As part of the 'Projet D'amélioration de la Qualité de l'Education' (Paque), the Ministry of Education has committed to a substantial expansion of RBF, encompassing all primary schools in 12 out of 25 provinces by 2022.

outcomes in particular. First, we revisit the debate about whether textbooks can increase student achievement. Prior randomized evaluations have primarily focused on understanding the value of having textbooks in schools to promote student learning (Glewwe et al., 2009; Sabarwal et al., 2014; Milligan et al., 2017), without special regard to incentives, and found no effect on average student performance. Consistent with recent studies that have underlined the important combined effect of providing inputs (like grants or textbooks) and incentives on educational outcomes (Mbiti et al., 2019; Gilligan et al., 2018), we study the impacts of a 'textbook for home-study' routine that combines incentives and textbooks, on student achievement. Prior work has also mainly focused on the use of textbooks within the classroom, whereas we explore the power of using textbooks at home to improve self-study and thus foster self-learning. In that sense our paper also ties into the literature on student motivation, self-efficacy, and self-learning. Self-study, defined as study outside the classroom and without direct supervision, is widely recognized as a valuable way to learn. To the best of our knowledge, our study provides the first causal evidence of the positive effects of self-study on student achievement in fragile and conflict-affected environments. One of the main challenges for students to self-study is to improve self-control and delay instant gratification (Beland and Murphy, 2016). In our setting, the textbook may also have served as a useful reminder or commitment device for students to follow-through on their intentions to self-study.

The rest of this paper is organized as follows. Section 2 describes the context and the intervention. Section 3 describes our data. Section 4 presents our main results. Section 5 discusses some related outcomes and the intervention's cost-effectiveness. Section 6 concludes.

2 Context and Intervention Description

2.1 Context

The Democratic Republic of Congo is one of the poorest, most conflict-ridden countries in the world. More than 80 percent of its population lives in extreme poverty. The country counts about 5 percent of the world's extreme poor and this percentage share is expected to double by 2030 (OECD, 2018). DRC is slowly recovering from a conflict known as Africa's 'First World War,' which led to the

loss of some five million lives between 1994 and 2003, and from subsequent conflicts in its eastern provinces. Among them is South Kivu. In 2015, around 60 percent of households were living below the national poverty line, putting the province roughly in the middle of the distribution of DRC provinces (IMF, 2015; Marivoet and De Herdt, 2015). Although this poverty rate is reportedly declining since 2005, daily lives in South Kivu remain marked by tension, and outbursts of violence carried out by various armed groups, including the regular army, are not infrequent (Kivu Security Tracer, 2018).

Despite recent efforts to improve the budget allocation, public finance for primary education, and education in general, remains low compared to most other countries in the region, with only 10.9 percent of the public budget allocated to education and with education budget execution at about 1.8 percent of GDP. The state budget supports two types of schools: the 'écoles conventionnées' or schools managed by the country's various religious networks and the 'écoles non-conventionnées' or regular public schools, managed and operated by government. The conventionné schools account for the majority of the country's publicly financed schools. Government spending is widely insufficient. State funds are spent mostly on salaries, and to a much smaller extent on the purchase of goods and services (UNESCO, 2014). Yet about two thirds of teachers nationwide are not on the official payroll. For their salaries, they rely heavily on a wide range of school fees, which are prohibitively costly for many poor households. Furthermore, public spending on education is uneven, biased towards the rich (World Bank Group, 2015), particularly in preschool, secondary and higher education. Even though South Kivu counts 15 percent of the country's primary-level students, the province receives only 7 percent of the teachers' budgets (De Herdt and Titeca, 2016). Aid money plays an important role to help fill these financing gaps.

Primary school education lasts six years and is compulsory for 6 to 11-year old children. Since 2005, primary school enrolment has improved considerably (with net enrolment increasing from 51 percent in 2005 to 79 percent in 2014), especially enrolment of girls and children in rural areas. Yet important challenges remain. Dropout and repetition rates are high, with the repetition rate hovering at over 10 percent (World Bank Group, 2015). Schooling is frequently interrupted not just due to outbursts of violence, but also because many families struggle to pay the school fees. Quality of education is also

lagging behind. According to the PASEC (a joint program of francophone countries to evaluate the education sector in member countries) assessment of primary school education in 2014, most grade 6 students in DRC were not sufficiently competent in reading or mathematics (WDR, 2018).

Many primary school teachers lack the requisite teacher qualifications or training and frequently also strong motivation. Teachers' wages are very low and their disbursement highly erratic, the result of commonly delayed payment by government and parents' irregular payment of school fees. Further, in schools where a results-based financing scheme exists, school personnel has strong financial incentives to secure and retain a high student enrolment rate. Consequently, teachers frequently spend considerable time away from class teaching during school hours to "chase children" back to school. These issues clearly call for important, albeit costly and difficult to implement, structural changes.

As part of a US$150 million World Bank grant aimed at improving public service delivery, the DRC government launched a nationwide initiative in 2008 to distribute 18 million textbooks in primary schools. This initiative was followed by another US$100 million grant, between 2012 and 2017, in which more than 22 million textbooks were distributed. An internal review (World Bank, 2017) reported that 93 percent of these books actually reached schools, but also highlighted that additional effort is needed in the future in order to ensure the efficient and continuous use of the delivered textbooks. In 2018, the DRC government, again with the support of the World Bank, launched a new 5-year program called 'Projet d'Amélioration de la Qualité de l'Rducation' or PAQUE, specifically focused on tackling the learning crisis in primary schools in 12 of its 25 provinces. This new initiative explicitly recognizes a need for schools to make better use of textbooks, both in French and in local languages (World Bank, 2017). As part of this new initiative, the DRC government is also introducing Results-Based Financing (RBF) in 1,350 primary schools located in the 12 focal provinces and has contracted Cordaid to provide technical assistance to this effect. Two of the RBF indicators that have been newly introduced relate to the use of textbooks in class and at home and were directly inspired by this textbook routine. Taken together, recent strategic priorities in DRC's education sector render the findings of our study of immediate policy interest.

2.2 The intervention

The 'textbooks for self-study' routine was implemented during five consecutive trimesters, from second trimester in school year 2016-2017 through the end of school year 2017-2018, in 45 randomly selected primary schools in the districts of Shabunda and Walungu, in South Kivu. Together with 45 schools in the control group, these schools represent the full population of primary schools where Cordaid had introduced a results-based financing scheme starting in 2008. The benefit of testing the impacts of the 'textbooks for self-study' routine at these schools was twofold. First, as part of the RBF scheme, there was already a monitoring system in place (with school visits on a quarterly basis), which we could extend to include monitoring of compliance with our own intervention. Second, we were able to integrate our extrinsic monetary incentives with the financial transfers paid to schools as part of the RBF scheme. This allowed us to economize on implementation costs, a point which we further discuss in the subsection on cost-effectiveness.

The routine was developed with the inputs of multiple staff members at Cordaid based both at headquarters in Den Hague and locally in Bukavu; primary school teachers of both conventionnées and non-conventionnées schools (through a series of focus group discussions); and officials of SECOPE ('Service for the control of the teacher payroll'), Provincial head of Education, and the Ministry of Education. Thanks to these diverse inputs, the routine succeeded in leveraging an already existing, widely available resource (instead of requiring new resources), fitting well with existing practices of homework (children attend half days in school) and in-class evaluations, and making use of an appropriate bundle of incentives. In the past, many development innovations or initiatives – take the US$100 million initiative to distribute textbooks across schools in DRC– have failed to realize their intended impacts largely because they had overlooked or underinvested in incentives to take-up or utilize them.

The intervention consisted of an incentivized scheme to encourage students to bring home textbooks on a regular basis. At each treatment school, a common set of initiatives were staged to help ensure clarity on the goals and practical details of the intervention by all relevant stakeholders. For instance, a team of facilitators, which we recruited and trained ourselves, held introductory meetings

with the headmaster and 5th grade teachers, with the parent committee and in class with 5th grade students during which children made a poster together about the initiative (which then was hung up in the classroom as a useful reminder). Further, a handful of questions were added to the standard survey that RBF-auditors conduct on a quarterly basis, specifically to assess the implementation of the new routine.

To encourage students, teachers and schools to stick with the intervention, a 'multi-level, multi-motive' incentive scheme was introduced, though with a main focus on the students. For students, the reward comprised of two elements. The first was an intrinsic, non-material individual reward based on a publicly announced (posters in the classroom) star system. A student earned a star each week that the student had: i) Taken home and returned in good condition her/his math and French textbooks according to the assignment. ii) Taken part in the weekly quiz. Similar public star systems have been found effective in other contexts (Ashraf, Bandiera and Jack, 2014). Note that the individual stars are not associated with any material benefits, the incentives work only through mechanisms of social status. The second was an extrinsic, material group reward consisting of handouts of school material such as notebooks and pens and pencils. This reward was meant to be given to classes in which the ratio of actual stars to possible stars (if all class students got stars every week) was at least 75 percent. We thus attempted to tap multiple motives, intrinsic (reputational or image; in-group favoritism) as well as extrinsic (in-kind), assuming these motives are complementary. In practice, however, the material benefit was in the end delivered to all classes irrespective of achievement. As this took place ex post and should have been a surprise to non-performing classes, it should not have impacted the perceived incentives.

For schools, the reward consisted of a small flat compensation, US$120 per year, for taking part in the experiment. With this compensation, we sought to address the schools' concern about missing and damaged books. This compensation was first to be used to buy new books to compensate for lost and damaged books. If there was money left after having compensated for the missing books, then the school could use that money to cover its general expenses. Schools (headmasters and teachers) were thus encouraged to help their students take care of the books and minimize any loss or damage of books

while keeping the program on track. Note that contrary to most incentive systems evaluated in the literature, our intervention focuses on the students, not the teachers. This is natural as the intervention emphasized self-study, but it also has the advantage of not creating (perverse) incentives towards teaching only what is being directly measured and/or lowering the bar of required student achievement.

3 Data and Study Design

3.1 Data

Our sample is comprised of 90 primary schools in the South Kivu region, where Cordaid had introduced a results-based financing scheme.[9] We stratified the sample by district, and then randomly assigned schools in each stratum to treatment or control group. We gathered detailed information about these schools and especially their 5th and 6th grade students making use of our own base- and end-line surveys combined with four primary data sources. By combining different data sources, we can mitigate concerns about common method bias, triangulate our data, and validate our main results.

(1) Surveys and Student Tests

We held multiple surveys at each of the 90 schools both at baseline, before the intervention, and at end-line, after 18 months' implementation of the intervention. In each school, we surveyed the headmaster and the 5th (baseline) and 6th (end line) grade teachers. We asked about their socio-economic status, their professional experiences, human capital, school climate, headmaster's leadership style, and for teachers, about their teaching efficacy and practices, and whether they regularly received feedback. These survey questions were mostly drawn from established survey instruments, in particular from the

[9] RBF schools were selected following three key eligibility criteria: (i) the school needed to have a minimum 'viable' size (computed as 26 students multiplied by the number of classes in the school; (ii) the school needed to be reachable/accessible by car; and (iii) the school needed to be gender mixed. Specifically, in Shabunda, the schools needed to be located within a 30 km radius of the center of Shabunda. To further narrow down the list of potential schools, Cordaid sought to ensure that the proportion of different types of schools (such as Catholic, Protestant, or non-religious public) in the sample reflected the true proportions in the population at large in each district.

Harvard Graduate School of Education and PASEC.[10] All questionnaires existed in French and the two most commonly used local languages, Swahili and Mashi.

At baseline, all present 5th grade students completed a survey in class. This survey included questions about the student's household situation, school-related attitudes and habits, whether the student works outside school hours, and aspirations for the future. Our enumerators also instructed students to fill in an in-class test, which covered both mathematics and French, mostly based off questions from PASEC (2014). The test (provided in the Appendix) covered basic grammar and reading comprehension in the French part, and basic numeracy (sequencing, basic arithmetic, fractions) in the math part. Questions are multiple choice, and there are overall eight questions in each subject, with a few sub-questions each. Students were given the time needed to complete the tests. The results provide us with a useful proxy measure of knowledge at baseline and help us control for potential non-observable heterogeneity among the students.

In each class, we randomly selected 12 students whose primary caretaker (with prime responsibility for the child's education) was also interviewed. We made sure that the gender composition of the 12 selected students was similar to that of the class as a whole. All parent surveys were conducted by community-based organizations (CBOs) using tablets. These CBOs had experience surveying households in the context of RBF. The parent questionnaire asked about the family's socio-economic situation, housing situation, and their involvement and communication with their child's primary school.

We repeated all the surveys at end-line, though this time interviewing the 6th grade teachers. Given the overall poor performance on the test at baseline, we added a few simpler questions to the end-line test. We also administered the same test to the teachers. At treatment schools, some extra survey questions, explicitly focusing on the intervention, were added.

[10] The HGS survey instruments are explained here: https://www.panoramaed.com/panorama-student-survey. PASEC is an international organization with the mission to evaluate education in the French speaking parts of the developing world.

(2) National Exam Data

We hand-collected the available performance-outcomes on the national exam, or 'Test National de fin d'Etudes Primaires' (TENAFEP). This high-stake test is taken at the end of 6th grade and certifies completion of primary education and is a requirement for students wishing to pursue secondary education. The TENAFEP comprises three subjects, all administered in French: French language, general knowledge and mathematics, and takes 60 minutes to complete. For each school, the Provincial Inspection of Education publishes how many students took the test, of those how many failed (all split by gender) and then the names and overall score only for those students who passed the exam. We matched these student names with the names from our student database using a level of tolerance for incorrect or alternative spelling of names and double-checked this matching manually. All students who were surveyed at baseline and end line and passed the test were matched.

(3) Basic School Characteristics

We hand-collected several basic statistics about each school from the archival records maintained by the Provincial Division of Education. This allowed us to gather information about the school size, the type of school (and if conventionnée, also its religious denomination), and the type of basic school infrastructure.

(4) School Performance Data

The results-based finance scheme requires auditors to collect information on a host of management practices at the school, including the quality of administrative and financial management, frequency and number of supervisory meetings, quality of teaching practice, continuity of the program and use of standardized teaching modules, and the extent to which the school is a learning organization. We construct three indices that capture the quality of the teachers, schools' climate, and principal's leadership. Table A.1 provides an overview of all variable descriptions.

3.2 Attrition and Randomization Check

Whilst stratifying allowed us to ensure balance between the treatment and control groups in terms of location (school district), we still confirmed, using baseline survey data, that balanced

randomization was successfully achieved on other key observable characteristics. We regressed each observable variable on the intervention dummy variable, with the constant indicating the mean value within control schools. As shown in Table 1, across the 35 tests, four variables -namely students' age, gender, hours worked after school, and frequency of eating breakfast- are significantly different from zero, though the imbalance is slight.[11] Further, it is a priori unclear what their joint influence on learning would be. Our randomization checks thus fairly reassuringly indicate that randomization was effective at generating samples that were balanced on schools' and students' observable characteristics, and we also present analysis with controls for those four variables that were slightly imbalanced in our regression tables.

From the baseline survey, 1,486 students were matched at end-line, whereas 1,332 were not matched. Attrition was thus very high. This reflects in large part the notoriously high dropout and repetition rates in rural DRC (World Bank, 2015). Further, the problem of attrition may have been exacerbated by the 'late' implementation of our end-line survey: due to concerns about safety end-line surveys were conducted when the school year was just about to come to a close, so some students were simply not present the day of the survey and the test. The attrition rate in the treatment and control schools were not significantly different (respectively, 47 and 50 percent), and the correlates of attrition were theoretically sound. Table A2 shows that, overall, older students, boys, students who work more outside school, and who attend school without organized study time at school were more likely to drop out of our sample. We also ran a specification interacting all variables with the treatment variable to check for differences in attrition predictors in treatment and control schools, of which two are significant –the size of the grade 5 class and the school's infrastructure. The sizes of those differences are, however, small. This suggests that although the high attrition rates may be a concern, except for the two mentioned variables, there is no systematic difference between the treatment and control schools in terms of the observable characteristics of the students missing. Nevertheless, given the high level of

[11] These results are confirmed when regressing all observable variables jointly against the intervention dummy. When excluding the five variables that were found to be significant above, an F-test of joint significance fails to reject the null of no difference between the treatment and control samples. Including the five variables, the F-test rejects the null hypothesis.

attrition we also present results with inverse probability weights to correct for attrition based on observables.

3.3 Compliance

Thanks to the termly field visits by the RBF verification agents of the Agence d'Achat de Performance (AAP), we were able to generate fine-grained measures of schools' compliance with the new routine.[12] Compliance was high overall, though weaker in seven schools in the first year of the intervention (see Table A3). We exploit the variance in compliance intensity to strengthen the internal validity of our main results and show below both intention-to-treat (ITT) and local average treatment effects (LATE).

In addition, we also sought to assess compliance *post hoc* using questions that we added to the end-line student, teacher, and parent surveys. More specifically, we asked about students' use of textbooks at home. We find that 81% of surveyed students in treatment schools reported having taken home a textbook in the past month, versus 39% of surveyed students in control schools. These data allow us to triangulate the compliance/verification data collected by AAP. Further, they lend support to the notion that the new routine effectively led students in treatment schools to make greater use of textbooks at home.

4 Results

In this section, we first explore how the 'textbook for self-study' routine affected the students' results on the French and math tests we designed. We then turn to the sit and pass rates at the national exam (TENAFEP), as well as the national exam scores available only for the students with a passing grade.

In the Appendix (Figures A1), we present the frequency distribution of student test scores, separately for the treatment and control schools and at base- and end-line. Student achievement in both

[12] Specifically, the following six dimensions were carefully evaluated: Whether (i) textbooks were returned in good state; (ii) a 'take home' log with student names and date was maintained; (iii) the star system was correctly used; (v) posters were hanging in the classroom; (vi) textbooks were taken home on a weekly basis; (vi) all project documentation was neatly organized. In addition, RBF verification agents also noted the number of weekly classroom tests and verified whether at least 75% of all students had received a star on a regular basis. Taken together, these verification data allowed us to gain detailed understanding of overall compliance, and 'early' detect (and attempt to remediate) schools that were experiencing difficulty.

control and treatment schools improved substantially between 5th and 6th grade, in particular at the lower end of the distribution. The 6th grade is a key year as it is the year of the national exam required to continue to secondary school, so the improvement is expected. The large size of the improvement still suggests that it is possible that the baseline test was conducted with less sense of importance than the end-line test, even though there is no clear rationale for why this should be the case. The distributions of baseline results look very similar in the control and treatment schools, though, and there is no statistically significant difference in unconditional average baseline test scores across the two groups. For end-line test results, the distributions look more different for French language, and the unconditional average score in the treatment group is significantly higher than in the control group in this subject. It is also important to note that there is a strong positive correlation between baseline and end-line test scores at the individual level, suggesting that also the baseline tests results are meaningful despite lower than expected test results.

4.1 French and Math Test Results

Following our pre-analysis plan[13], we estimate the effects of the 'textbook for self-study' routine using the following linear model

$$Y_{i,j,d,s,t+1} = \alpha + \beta Treatment_j + \pi Y_{i,j,d,s,t} + \gamma \phi_{idt} + \epsilon_{i,j,s,t+1} \ , \qquad (1)$$

where Y_{ijdst} is student i's standardized test score in school j in district d in subject s at time t; *Treatment* is the indicator variable for being in a treatment school; and ϕ is a vector of district and enumerator fixed effects.[14] Each enumerator were assigned both treatment and control schools and ran over 100 student tests on average, allowing us to remove enumerator fixed effects from all empirical specifications. This helps to address the concern that enumerators may have induced differential measurement error, for instance by helping students comprehend questions on the test or influencing

[13] The analysis follows the pre-analysis plan published on May 8th, 2018, at the AEA RCT Registry. The RCT ID is AEARCTR-0001845.

[14] In principal randomization should take care of the well-known challenge that knowledge is a cumulative process that depends on the full history of a subject's exposure to educational input, not just a time limited recent intervention. Including baseline test results as sufficient statistics for prior input into learning should thus not be as essential in this case (Todd and Wolpin 2003). Having baseline results, though, helps us reduce noise from unobserved heterogeneity within our sample.

the way the intervention was grounded from the start (West and Blom, 2017; Crossley et al., 2017). For every specification, we confirmed that the coefficients on the enumerator dummies were indeed significantly different from zero. Error terms are clustered at the school level to take into consideration intra-cluster correlations. In an alternative specification, we additionally include the four variables that appeared slightly unbalanced before the intervention.[15]

As shown in Table 2a, we find that students who were in the treatment schools scored 0.281σ to 0.319σ higher in French compared to students in the control group schools (Columns 1-3), but scored only marginally, 0.022σ to 0.036σ, and not significantly higher in math (Column 4-6). For both French and math, we first report the base model regression results, then the estimates when controlling for the four variables found to differ between treatment and control schools at baseline, and finally additionally including the inverse probability weights to adjust for potential attrition bias. Though our results are slightly stronger both in terms of economic and statistical significance when we adjust for potential attribution bias, our major findings hold across all three empirical specifications. This is reassuring as the relatively high level of attrition could potentially have affected the results.

As discussed above, in seven of the treatment schools, the intervention was not properly implemented in the first year. In Table 2b we therefore provide the local average treatment effect (LATE) instrumenting actual treatment with intended treatment status. As expected, the estimated impact goes up, for French to 0.359σ. There is, however, still no significant difference found in math scores.[16] Estimated coefficients in different models are also illustrated in Figure 1.

In addition to presenting impacts on standardized total scores, we also present impacts on different domains of subject-level competencies in Figure 2 using the ITT model with inverse probability weights. The effects are positive and significant across the two domains covered by our test of French language competencies (sentence completion and retrieving explicitly provided information).

[15] We replaced the missing observations for the teaching efficacy variable with the mean value (in the district) and added a dummy variable to control for observations imputed in this fashion.
[16] It is important to note that our baseline test shows no significant differences between complying and non-complying schools, but we also have a small sample (only 7 schools, 108 students, are non-complying).

The effects on results across the three math-specific domains (decimals, sequencing and word-based math problems) are non-significant. Still, results indicate a weak improvement in performance on word based math problem, which could be linked to the found improvement in French test score results.

Drawing on prior research, we highlight several plausible explanations for why the textbook routine produced a positive impact on students' achievement in French and not in math. First, this discrepancy may well reflect known differences in the knowledge accumulation process between these two subjects (de Jong, 2015). In math, the knowledge elements build upon each other, whereas in French, the knowledge elements are more parallel. Therefore, when the overwhelming majority of students do not pass a proficiency threshold in math (World Development Report, 2018) or lack understanding of the very basic math concepts, the marginal benefit of using a grade-appropriate (read, a more advanced) math textbook may be especially small. Second, for students to make measurable progress in math, the textbook routine may have been insufficient. Indeed, prior work has argued that more help and supervision in school and/or at home is typically needed for students to progress in math (Lee and Barro, 2001; Marcotte, 2007). Third, this discrepancy may well reflect differences in the textbooks' readability. Language textbooks (French in the case of DRC) tend to be written at a more rudimentary level than subject textbooks (Chimombo 1989; Milligan et al. 2016), and thus also at a more appropriate level for students that lag behind. Since math textbooks were written in the French language, there was also a double treatment of French language exposure. Note since our intervention obliged all students to take home both a math and French textbook at least once per week, we can refute that differential exposure to these textbooks drives the discrepancy.

4.2 TENAFEP

Turning to the high-stake national exam, we estimate an equation of similar form as equation (1) to first evaluate the effects of the intervention on students' likelihood of passing the TENAFEP (Table 3, Columns 1-2). We find that the intervention increased the average likelihood of successfully passing the exam in treatment schools by about 10 percent. This is encouraging, as passing the national exam is a requirement for continuing to secondary school.

One limitation though of the student-level data is that they do not allow us to identify to what extent the increased probability of TENAFEP success is due to an increase in the number of students taking the exam versus an increase in the success rate of these students. To address this question, we combined the TENAFEP data with school-level administrative data on the number of students in 6th grade. We are thus able to separately estimate the likelihood that a student took the TENAFEP exam and the likelihood of success at school-level. Results, presented in Table 4, show that the intervention significantly raised the number of students taking the exam, but not the actual pass rate. Put differently, even though the intervention led more students to take the exam, the average TENAFEP pass rate did not drop; rather, it remained the same. Importantly, the size of the treatment effect on TENAFEP sitters is of similar magnitude to the treatment effect on TENAFEP passers (Columns 5 and 6).

Turning to the TENAFEP scores, it is important to keep in mind that the increased number of students sitting the exam is likely to influence the average score. If the additional students disproportionally are drawn from the pool of weaker students, which is quite likely, then the average score will go down. It should also be noted that the TENAFEP consists of three parts (math, French and "other") but only the aggregate test results are made available. Nevertheless, as shown in Table 3, Columns 3 and 4, the estimated effect of the treatment dummy on the TENAFEP z-score is positive, and equal to 0.19σ, though it is not statistically significant.

In sum, the intervention led more students to sit and pass the high stakes national exam, which is very encouraging as having passed the TENAFEP is a requirement to continue onto secondary school education. Whilst it appears that the earlier evidenced gains in French were insufficient to significantly raise the overall average TENAFEP score, it is noteworthy that the average score did not drop, given that additional students taking the test were most likely drawn from the lower part of the student performance distribution.

5 Discussion

5.1 Habits, Attitudes and Aspirations

In this section we provide some additional evidence to shed light on the impact of the 'textbook for home-study routine' on students' habits and aspirations related to learning and school performance. We estimate an equation similar to (1) with inverse probability weights to adjust for attrition bias, and with as dependent variable a student's attitudes towards homework and textbooks, but also motivation to go to school, career aspirations and attitudes towards school and teachers, more generally. We control for baseline values of these survey measures, as well as the four student characteristics found significantly different between control and treatment groups in the balance table. Standard errors are also adjusted for multiple hypothesis testing. Table 5 reveals that students in the treatment schools if anything spent less time on homework, thereby ruling out the possibility that the impact of the new routine simply came from a higher homework load. Students in the treatment schools found the textbooks significantly more useful for learning than students in the control schools. Furthermore, they were almost 14 percent more likely to aspire to a non-manual job and 6 percent more likely to say that they wanted to continue school after primary (school ceases to be compulsory after 11 years old in DRC). Interestingly, the intervention had no measurable impact on (self-reported) in-class interaction, learning in class or motivation to go to school. These results suggest that the main mechanisms through which better learning outcomes were achieved operated at the individual student-level. Thanks to a more positive attitude towards textbooks, and more ambitious job and study aspirations, the intervention succeeded not only in raising student achievement but also encouraging more students to take the TENAFEP, a prerequisite for most non-manual jobs and secondary school.

5.2 Effect Size and Cost Effectiveness

The ITT effect of 0.281 to 0.319σ that we find for French language falls in the middle of the distribution of estimated effect sizes presented in the overview by Kremer et al. (2013). Glewwe et al. (2009) and Sabarwal et al. (2014) evaluate textbook allocation projects in Kenya and Sierra Leone, respectively, but neither finds a significant impact of textbooks on average student learning. For stronger students, though, Glewwe et al. (2009) estimate an impact of 0.22σ for the 5th quintile and

0.14σ for the fourth quintile after one year of exposure.[17] It should be noted that our intervention measures what can be thought of as the "intensive margin" of school inputs, making more use of existing textbooks, not the extensive margin, the impact of providing more books. We have not been able to identify directly comparable impact evaluations in the literature.

Cost effectiveness, and even low absolute costs, is particularly relevant in very poor and fragile settings such as DRC. Based on operating expenditures, we estimate the cost per student to US$17 in our treatment subsample. This includes the direct costs for the incentives to schools (roughly US$9 per student), i.e. school material shared with students over the school year and a flat compensation of US$120 per school. It also includes the costs of setting up the intervention and monitoring of compliance, including primarily costs of manpower and project management but also some small expenditures on survey material.

Based on our most conservative estimate of impact (0.281), this suggests that US$100 yields 1.6σ improvement in test scores, or alternatively that US$63 would achieve a 1 standard deviation improvement. This compares very favorably with the 30 RCTs evaluated for cost efficiency in Kremer et al. (2013). A significant part of the costs of this intervention are associated with the need to monitor schools for compliance and distributing financial compensation. The planned scaling up of RBF to 12 provinces covering 1 350 primary schools is hugely helpful in this case as the textbook routine could be embedded in that broader system of monitoring and incentives, and it should be possible to reduce costs per school. This being a bundled intervention, it is difficult to separate the roles of the financial incentives versus the non-pecuniary incentives through the star-system, but the direct costs of compensation to the schools for the financial incentives were slightly more than half of total costs. If

[17] Other studies have analyzed alternative pedagogical tools, such as flip charts (Glewwe et al., 2004) and multi-level learning materials (Tan et al., 1999). While the former finds no significant impact, the latter finds a very high impact on English in a Filipino context of a combination of multi-level learning materials and enforced parent-teacher partnerships (between 0.75 to 1.05 standard deviations). Recently, technology driven interventions geared towards teaching at the right level has shown potential in poor but stable environments. Banerjee et al. (2007) found that a computer aided learning program that provided two hours per week of math instruction improved test scores by 0.48 standard deviations after two years. Muralidharan et al. (2018) estimate ITT effects of 0.37 standard deviations in math and 0.23 in Hindi over a 4.5-month period from a personalized technology-aided after-school instruction program in urban India. Implementing something similar in a fragile rural setting such as Eastern DRC would be very challenging.

scaled up, reducing financial compensation and stressing more non-pecuniary incentives together with reduced monitoring costs through the RBF system, could potentially almost double cost efficiency. This is assuming, though, that the impact carries over also if scaled up, which of course cannot be guaranteed (e.g. Bold et al. 2018).

The compliance data also suggests that initial concerns about books disappearing were not well founded. The average (median) number of textbooks that disappeared or were damaged over the course of the intervention was, respectively 6.71 (3) in the first semester of the intervention and 5.3 (3.5) in the second semester.

5.3 Heterogenous Treatment Effects

The main results in Section 4.1 address the impact of the intervention on average test score results. In this section we turn to analyze the impact across different groups of students and teachers. This analysis should be interpreted with some care, as we are restricted in sample size and therefore statistical power. Differences in impact would have to be quite substantial to be deemed statistically significant, so we cannot for sure rule out differences of smaller size. All tables are presented in the appendix.

In Table A4 we investigate whether impact vary with the status of the students. It is possible that for instance academically stronger students, students from better home conditions and with less responsibilities at home (something that likely varies with gender) have greater opportunities to benefit from the intervention. We therefore study variation by gender, status as vulnerable/indigent,[18] age (older students have typically repeated grades), and baseline test scores using a linear interaction in the model with inverse probability weights. For math, neither the linear effects nor the interaction effects are statistically significant in any of the models. For French, we find no significant difference depending on gender or age. We do find, though, that students classified as vulnerable did not benefit from the intervention, while the academically weaker half of the students at baseline benefitted more than the

[18] Indigents (particularly vulnerable children) are formally defined as children who are orphaned by father or mother or both, HIV / AIDS orphans, with physical disabilities, whose parents are identified as vulnerable due to physical or mental disability, or children who have been displaced by war.

stronger students. Viewed from an equity standpoint, the evidence suggests that the intervention did not unintentionally reinforce any gender or age discrimination or biases against the academically weakest students in class, if anything the opposite. At the same time, though, the intervention had no impact on the students with the most challenging home conditions.

Next, we investigate whether the impact vary with teacher characteristics, notably teacher's teaching efficacy using the Harvard Instrument, years of work experience, and teacher competence in math and French as measured by their score on the student math and French tests. Teacher competence has been shown to be low across Sub-Saharan Africa and to negatively impact student learning (Bold et al. 2017). Results in Table A5 tentatively suggest that the new routine if anything had a bigger impact on French test scores in schools with weaker teachers, in particular with regards to teaching experience. This should be interpreted with some caution, but the results are consistent with the notion that textbooks can somewhat compensate for weak in-class learning in the French language.

In sum, these results should be interpreted with some care, but tentatively suggest that students generally benefit in French language skills from the intervention, but the academically weaker students more so and the socio-economically most vulnerable less so. The results also tentatively suggest, though, that textbooks at home and self-learning may be particularly beneficial for students in schools with weaker teachers. Teacher competence and textbooks thus serve as substitutes more than complements in the production of learning when the focus lies on usage of textbooks for self-study.

6. Conclusion

The concept of a learning crisis in the developing world has turned attention away from school enrolment to student achievement. Numerous interventions have been implemented and evaluated, both at the demand and the supply side. A set of recent papers have emphasized the complementarity of incentives to stimulate effort and accountability, and school inputs such as textbooks (e.g. Mbiti et al. 2019; Gilligan et al. 2018), arguing that it takes the combination of the two to generate significant impact on student learning. This insight may matter the most in severely resource constrained and

fragile countries. Very few experiments have been properly evaluated in these settings even though this is where the main challenges looking forward lie and interventions that are effective in poor but stable environments may not be feasible or effective in fragile settings. Interventions that use incentives to make better use of existing resources are particularly useful given the shortages these countries are facing.

Over an 18-month period we implemented a novel incentivized textbook routine aimed at strengthening students' self-study at home and thereby make better use of existing textbooks in 45 primary schools in South Kivu, the Democratic Republic of Congo. The 45 schools were chosen randomly and compared with an equal number of control schools. We show evidence that the intervention led to 0.281 to 0.359 standard deviations increase in test scores in French, relative to the control group, but had no measurable effect on test scores in math. The intervention also raised the number of students receiving a passing grade at the national exam (TENAFEP), a high-stake test that is a requirement to be eligible for secondary school. This increase was due to more students sitting the test, and despite the increase in numbers, average test scores did not go down, if anything the opposite. That the intervention led more students to pass the exam is encouraging, as it is necessary for students to proceed to the next level of education.

The intervention seems to have operated mostly at the individual student-level. We found that the intervention led students to express a more positive attitude towards textbooks and more ambitious aspirations for future education and work. Even at this small scale, the intervention also compares favorably in terms of cost efficiency. Based on the most moderately estimated ITT effect of 0.28σ, we estimate that $US 100 yields a 1.6σ improvement in test scores. If implemented on a larger scale, some costs can be reduced, but there are also challenges in terms of monitoring and securing compliance.

To raise student learnings in rural DRC and other similar fragile settings to levels reflecting the ambitions and goals set in the official curriculum and global targets requires substantial improvements in most aspects relevant for knowledge production, from household level inputs at early age over teacher quality to school management. Such an effort requires substantial resources, both financial and human, and depends strongly on the ability to generate stability and inclusive growth. To start identifying cost

efficient interventions that largely rely on clearer incentives and more efficient use of existing resources can be a first step, though, towards making a difference in the life of students.

For scaling up, more experimentation can help to reach the full potential, and design the most cost-efficient implementation. Training teachers in the usage of textbooks both in the classroom and for self-study could potentially increase the impact of the routine. This could include using textbooks from different grades as a diagnostic tool to reach students at their individual level, thus avoiding some of the problems associated with overambitious curricula as highlighted in for instance Beatty and Pritchett (2012). Textbook language is also a salient challenge, in particular if the routine was to be expanded to topics requiring even more text analysis. The DRC government together with donors are currently distributing textbooks in four local languages in lower grades. Making sure these books are used to their potential and evaluating the role of language for both learning and ambitions for the future, could yield important insights. Finally, limited resources prevented us from analyzing properly the relative role of financial versus non-financial incentives for students and schools. For cost efficiency, understanding what primarily drives change in behavior, and how to secure compliance through monitoring, becomes essential if scaling up this intervention.

Our study raises intriguing questions for future research. What are other approaches (peer to peer learning) that sideline the challenge of poor teacher quality and weak parental support, and how can they be leveraged to improve learning in primary schools? How might the impact of a textbook intervention differ if one considered a younger cohort? How might a change in incentives impact outcomes? Will our main findings replicate in other settings? We defer these questions to further study.

References

Ashraf, N., Bandiera, O. and Jack, B.K., 2014. “No margin, no mission? A field experiment on incentives for public service delivery.” *Journal of Public Economics*, 120, pp.1-17.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden, 2007. “Remedying Education: Evidence from Two Randomized Experiments in India.” *Quarterly Journal of Economics*, 1235-1264.

Barro, Robert and Jong Wha Lee, 2013. “A new data set of educational attainment in the world, 1950–2010.” *Journal of Development Economics*, 104, pp. 184-198.

Beatty, Amanda and Lant Pritchett, 2012. "The Negative Consequences of Overambitious Curricula in Developing Countries," Scholarly Articles 9403174, Harvard Kennedy School of Government.

Behrman, Jere R., Susan W. Parker, Petra E. Todd, and Kenneth I. Wolpin, 2015. "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools." *Journal of Political Economy* 123(2): 325-364.

Beland, Louis-Philippe, and Richard Murphy, 2016. “Ill Communication: Technology, Distraction & Student.” *Labour Economics* 41: 61-76.

Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Brian Stacy, Christophe Rockmore, Jakob Svensson, 2017. “Enrollment without Learning: Teacher Effort, Knowledge, and Skill in Primary Schools in Africa.” Journal of Economic Perspectives, 31, pp. 185–204.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabub, Alice Ng’anga, Justin Sandefur, 2018. “Experimental evidence on scaling up education reforms in Kenya.” *Journal of Public Economics*, 168, 1-20.

Burde, Dana and Leigh L. Linden, 2013. “Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools.” *American Economic Journal: Applied Economics*, 5(3), 27-40.

Burde, D., Kapit, A., Wahl, R.L., Guven, O. and Skarpeteig, M.I., 2017. "Education in emergencies: A review of theory and research. *Review of Educational Research*, *87*(3), pp.619-658.

Chimombo, M., 1989. "Readability of Subject Texts: Implications for ESL Teaching in Africa." *English for Specific Purposes* 8 (3): 255–264.

Crossley, Thomas, Tobias Schmidt, Panagiota Tzamourani and Joachim K. Winter, 2017. "Interviewer Effects and the Measurement of Financial Literacy", ISER WPS No. 2017-06, University of Essex.

De Herdt, Tom and Kristof Titeca, 2016. "Governance with Empty Pockets: The Education Sector in the Democratic Republic of Congo", *Development and Change*, Vol. 47 (3): 472–494.

de Jong, John H. A. L. 2015. "Why Learning English Differs from Learning Math and Science." Pearson, October 20. Available on https://www.english.com/blog/learning-english-differs-from-learning-maths-science-gse/.

Fredriksen, Birger, Sukhdeep Brar and Michael Trucano, 2015. "Getting Textbooks to Every Child in Sub-Saharan Africa : Strategies for Addressing the High Cost and Low Availability Problem." Directions in Development--Human Development; Washington, DC: World Bank.

Gilligan, Daniel O., Naureen Karachiwalla, Ibrahim Kasirye, Adrienne M. Lucas, Derek Neal, 2018. "Educator Incentives and Educational Triage in Rural Primary Schools." NBER WP 24911.

Glewwe, Paul and Karthik Muralidharan, 2016. "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications", in *Handbook of the Economics of Education*, pp. 653-743. Elsevier.

Glewwe, Paul, Michael Kremer, Sylvie Moulin and Eric Zitzewitz, 2004. "Retrospective vs. Prospective Analyses of School Inputs: The Case of Flip Charts in Kenya". *Journal of Development Economics* 74(1):251-168.

Glewwe, Paul, Michael Kremer, and Sylvie Moulin, 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics*, 1(1): 112-35.

IMF, 2015. “Democratic republic of Congo: Selected issues”, IMF Country Report No. 15/281, Washington DC.

Jacob, Brian A., and Steven D. Levitt, 2003. “Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating.” *Quarterly Journal of Economics* 118 (3): 843–77.

Kremer, Michael, Conner Brannen, and Rachel Glennerster, 2013. “The Challenge of Education and Learning in the Developing World.” *Science* 340, 297-300.

Lee, Jong-Wha and Robert J. Barro, 2001. "Schooling Quality In A Cross-Section Of Countries." *Economica*, 68, 465-488.

Marcotte, D.E., 2007. “Schooling and test scores: A mother-natural experiment.” *Economics of Education Review*, 26, 629–640.

Marivoet, Wim and Tom De Herdt, 2015. “Poverty lines as context deflators: A method to account for regional diversity with application to the Democratic Republic of Congo”, *Review of Income and Wealth*, 61(2):329-352.

Martinelli, César, Susan W. Parker, Ana Cristina Pérez-Gea, and Rodimiro Rodrigo 2018. “Cheating and Incentives: Learning from a Policy Experiment.” *American Economic Journal: Economic Policy*, 10(1): 298–325.

Mbiti, Isaac , Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, Rakesh Rajani, 2019. “Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania.” NBER WP 24876.

Milligan, Lizzi O., John Clegg and Leon Tikly, 2016. “Exploring the potential for language supportive learning in English medium instruction: a Rwandan case study.” *Comparative Education*, 52:3, 328-342.

Milligan, Lizzi O., Leon Tikly, Timothy Williams, Jean-Marie Vianney, Alphonse Uworwabayeho, 2017. "Textbook availability and use in Rwandan basic education: A mixed-methods study." *International Journal of Educational Development*, 54, pp. 1-7.

Muralidharan, Karthik, Abhijeet Singh and Alejandro J. Ganimian, 2018. " Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India." *American Economic Review*, accepted September 2018.

OECD, 2015. "The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence." PISA, OECD Publishing, Paris.

Orkin, Kate, 2013. "The Effect of Lengthening the School Day on Children's Achievement in Ethiopia." YL-WP119.

PASEC, 2015. "PASEC2014: Education System Performance in Francophone Sub-Saharan Africa." PASEC, CONFEMEN Dakar.

Read, Tony, 2015. "Where have all the textbooks gone?: toward sustainable provision of teaching and learning materials in Sub-Saharan Africa." Directions in development; human development. Washington, D.C.: World Bank Group.

Sabarwal, Shwetlena, David K. Evans, and Anastasia Marshak, 2014. "The permanent input hypothesis: the case of textbooks and (no) student learning in Sierra Leone", Policy Research working paper, no. WPS 7021. Washington, DC: World Bank Group.

Tan, Jee-Peng, Julia Lane, and Gerard Lassibille, 1999. "Student outcomes in Philippine elementary schools: An evaluation of four experiments." *World Bank Economic Review*, 13 (3), 493-508.

Todd, Petra and Kenneth I. Wolpin, 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement," *The Economic Journal*, 113, pp. 3-33.

UNESCO, 2013. "The Global Learning Crisis: Why every child deserves a quality education", UNESCO, Paris.

UNESCO, 2014. “Education for All Global Monitoring Report.” UNESCO.

West, Brady T. and Annelies G. Blom 2017. “Explaining Interviewer Effects: A Research Synthesis”, *Journal of Survey Statistics and Methodology* 5, 175-211.

World Bank, 2017. “Implementation Completion and Results Report (TF-14253, TF-14358)”, Report No: ICR00004233, Washington DC: World Bank.

World Bank, 2018. “World Development Report 2018: Learning to Realize Education’s Promise”, Washington DC: World Bank.

FIGURES AND TABLES

Figure 1: ITT effect on student achievement for different model specifications

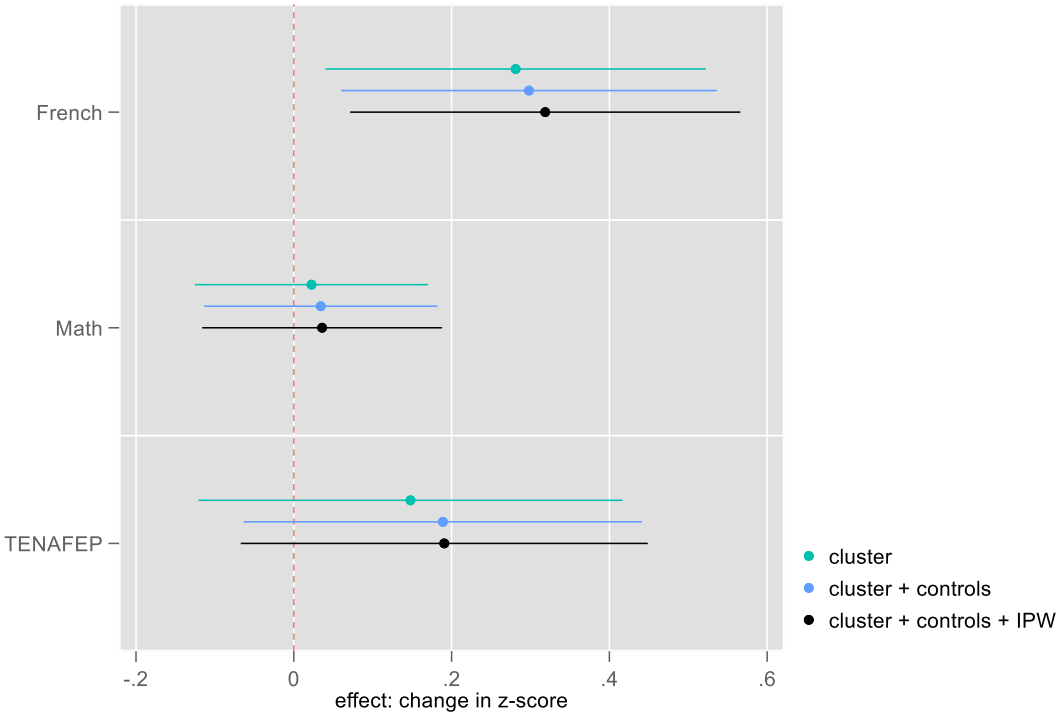


Figure 2: ITT effects by specific competence assessed (model is cluster + controls + IPW)

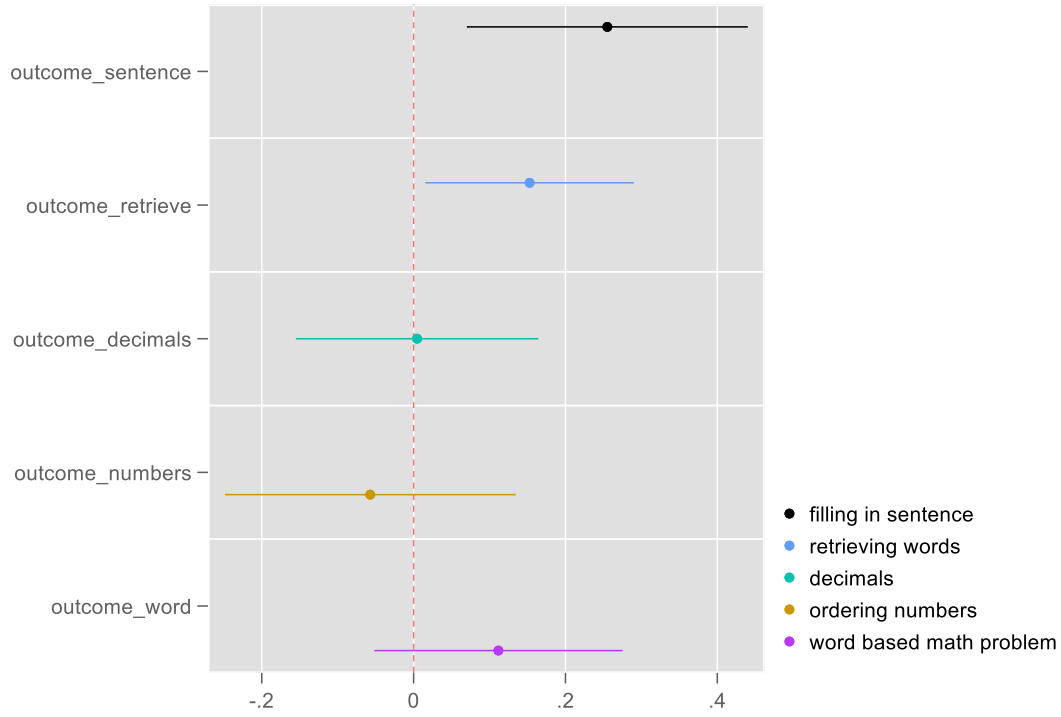


Table 1: Key characteristics at baseline

	control group			treatment group			(7)
	(1) mean	(2) SD	(3) n	(4) mean	(5) SD	(6) n	difference control – treatment
Pupil-level characteristics							
indigent (d)	0.112	0.316	704	0.111	0.314	794	-0.001
age (years)	11.844	1.356	704	12.077	1.344	794	0.233**
girl (d)	0.477	0.500	704	0.530	0.499	794	0.058*
mother literacy	0.628	0.484	704	0.651	0.477	794	0.023
father literacy	0.828	0.378	704	0.819	0.386	794	-0.009
support for homework	0.283	0.451	704	0.293	0.456	794	0.011
minutes on homework	40.653	26.709	704	42.469	28.918	794	1.815
hours of work (non-school) per week	5.766	3.073	704	4.915	3.107	794	-0.851*
experience of violence at school (0-4)	1.980	1.227	704	2.261	1.325	794	0.281
frequency of eating breakfast (0-3)	1.974	1.214	704	1.686	1.307	794	-0.288*
math score at baseline (/12)	3.330	3.210	704	3.185	3.014	794	-0.204
French score at baseline (/16)	2.638	2.501	704	2.433	2.316	794	-0.001
difference between baseline and endline math score (z-score)							0.124
difference between baseline and endline French score (z-score)							-0.529*
School-level characteristics							
Organised study time at school after school	0.370	0.488	46	0.477	0.505	44	0.143
Teacher supervises during this time	0.087	0.285	46	0.091	0.291	44	-0.004
Ratio success at TENAFEP	0.574	0.436	45	0.597	0.407	44	0.211
Ratio register at TENAFEP	0.678	0.326	45	0.691	0.306	44	0.004
Teachers per pupil	0.005	0.009	45	0.005	0.009	44	0.001
Size of grad 5 class	52.156	34.365	45	51.136	35.898	44	-2.030
Ratio vulnerable pupils	0.095	0.037	45	0.100	0.041	44	0.006
Ratio girl/boy	1.038	0.177	45	1.035	0.208	44	-0.010
Ratio of revenue from RBF	3.282	1.096	44	3.402	1.410	44	0.143

Number of brick walls	9.511	3.806	45	9.182	4.228	44	-0.437
Regular meeting with parents (d)	0.711	0.458	45	0.727	0.451	44	0.012
Mean test score	4.118	3.296	46	3.798	3.284	44	-0.628
Shabunda district	1.565	0.910	46	1.545	0.901	44	-0.073
RBF index on teacher performance	26.489	12.050	45	26.273	13.769	44	-0727
RBF index on quality	7.422	5.097	45	6.773	3.298	44	-0.870
School leadership (Harvard instrument 1-5)	3.833	0.337	45	3.788	0.288	44	0.036
School climate (Harvard instrument 1-5)	3.804	0.377	45	3.792	0.215	44	-0.011
Grade 5 teacher-level characteristics							
Years of work experience of teacher(s)	13.778	11.231	45	14.818	11.517	44	0.841
Teaching efficacy (Harvard instrument)	1.527	0.127	46	1.516	0.083	44	-0.009
Average score on French pupil test (maximum 27)	23.237	2.333	38	23.500	1.860	39	-0.396
Average score on Math pupil test (maximum 21)	13.778	11.231	45	14.818	11.517	44	0.185

Note: column 7: P-value * < 0.1 ; ** < 0.05 ; *** < 0.01 Clustered standard error at the school level.

Table 2a: Intent-to-treat Effects in an OLS regression framework

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Standardized test scores (end line)					
	French	French	French	Math	Math	Math
Intervention (ITT)	0.281*	0.298**	0.319**	0.022	0.034	0.036
	(0.146)	(0.145)	(0.150)	(0.090)	(0.090)	(0.092)
Baseline z-score	0.143***	0.140***	0.135***	0.219***	0.211***	0.211***
	(0.049)	(0.049)	(0.049)	(0.041)	(0.041)	(0.039)
Pupil controls (4)	No	Yes	Yes	No	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Inverse Probability Weights	No	No	Yes	No	No	Yes
Observations	1498	1498	1498	1498	1498	1498
Adjusted r-squared	0.298	0.299	0.288	0.286	0.294	0.281

Note: Robust standard errors clustered by school in parentheses. Treatment is a dummy variable whether students were attending a treatment school. Tests in French and math were designed to cover wide ranges of achievement and to be linked between baseline and end line assessments using common items. Scores were standardized to have a mean of zero and standard deviation of one in the baseline. P-value * < 0.1 ; ** < 0.05 ; *** < 0.01

Table 2b: LATE Effects in an OLS regression framework

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: Standardized test scores (end line)					
	French	French	French	Math	Math	Math
Intervention (ITT)	0.315**	0.336**	0.359**	0.026	0.038	0.040
	(0.159)	(0.158)	(0.165)	(0.098)	(0.098)	(0.101)
Baseline z-score	0.138***	0.135***	0.131***	0.218***	0.210***	0.210***
	(0.048)	(0.048)	(0.049)	(0.041)	(0.040)	(0.038)
Pupil controls (4)	No	Yes	Yes	No	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Inverse Probability Weights	No	No	Yes	No	No	Yes
Observations	1498	1498	1498	1498	1498	1498
Adjusted r-squared	0.305	0.307	0.296	0.287	0.295	0.283

Note: Robust standard errors clustered by school in parentheses. Treatment is a dummy variable whether students were attending a treatment school. Tests in French and math were designed to cover wide ranges of achievement and to be linked between baseline and end line assessments using common items. Scores were standardized to have a mean of zero and standard deviation of one in the baseline. P-value * < 0.1 ; ** < 0.05 ; *** < 0.01

Table 3: Effects on standardized test score (TENAFEP)

	(1) LPM TENAFEP	(2) LPM TENAFEP	(3) OLS TENAFEP z-score	(4) OLS TENAFEP z-score
Intervention (ITT)	0.097*	0.103**	0.141	0.190
	(0.050)	(0.047)	(0.164)	(0.155)
Control baseline test	Yes	Yes	Yes	Yes
School clustered SE	Yes	Yes	Yes	Yes
Pupil controls (4)	No	Yes	No	Yes
Enumerator controls (baseline only)	Yes	Yes	Yes	Yes
District control	Yes	Yes	Yes	Yes
N	1371	1371	806	806
R-squared	0.077	0.090	0.112	0.147

Note: Standard errors in parentheses. P-value * < 0.1 ; ** < 0.05 ; *** < 0.01 | X: for Logit models we are reporting the average marginal effect of benefitting from the intervention | district controls are included to account for the fact that TENAFEP score are weighted at district level (giving a bonus to Shabunda for its remoteness).

Table 4: Effects on standardized test score (TENAFEP) – school-level data

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
	Sit rate	Sit rate	Pass rate	Pass rate	TENAFEP (score /100)	TENAFEP (score /100)
Intervention school	0.094**	0.090*	-0.001	-0.004	1.142	1.095
	(0.046)	(0.046)	(0.022)	(0.022)	(0.909)	(0.957)
control baseline value	Yes	Yes	Yes	Yes	Yes	Yes
school-level control (1)	No	Yes	No	Yes	No	Yes
district control	Yes	Yes	Yes	Yes	Yes	Yes
N	83	83	83	83	82	82
R-squared	0.032	0.024	0.084	0.077	0.088	0.077

Note: Standard errors in parentheses. P-value * < 0.1 ; ** < 0.05 ; *** < 0.01 | No end line TENAFEP data was available for 7 schools (5 in the treatment group, 2 in the control group). | See Table 3 for an explanation for the inclusion of a district control.

Table 5: Habits, Attitudes and Aspirations

	(1) Help with Homework (binary)	(2) Time spent on homework (0-6)	(3) Finds manuals useful for learning (0-4)	(4) Aspires to non-manual jobs (binary)	Wishes to go to secondary school (binary)	(5) Motivation to go to class (0-4)	(6) Is learning in class (0-4)	(7) Asks questions (0-4)	(8) Uses manual in class (0-4)
Treatment	0.012	-0.198	0.401***	0.136***	0.062***	0.036	0.031	0.052	0.234
	(0.058)	(0.125)	(0.148)	(0.046)	(0.020)	(0.110)	(0.110)	(0.092)	(0.160)
Bonferroni p-value	1	0.522	0.048	0.059	0.028	1	1	1	0.751
Baseline test results	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Pupil controls (4)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Enumerator fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inverse Probability Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1498	1498	1497	1497	1497	1497	1398	1440	1391
Rsquared									

Note: A non-manual job is defined as a job belonging to one of the following categories: artist, lawyer, politician, faith minister, and civil servant –including teachers but excluding military and police. We do not have a baseline for pupil's professional aspirations, but given that randomization largely created balanced treatment and control groups this should not be a concern.

Appendix: Complementary Tables and Results

Table A.1: Description of variables assessed at baseline		
VARIABLES	Description	Data source
Organized study (d)	Indicator whether organized study time at school after school exists	Own survey
Accompanied study (d)	Indicator whether organized study time supervised by a teacher at school after school exists	Own survey
Ratio success at TENAFEP	Share of students who passed the of those who took the TENAFEP	RBF audit
Ratio register at TENAFEP	Share of students who registered for the TENAFEP of those enrolled in the sixth grade	RBF audit
Teachers per pupil	Ratio number of teachers over number of pupils	RBF audit
Size of grade 5 class	Number of students in the 5^{th} grade class	RBF audit
Ratio vulnerable pupils	Share of students with vulnerability status in school	RBF audit
Ratio girl/boy	Share of girls in school	RBF audit
Ratio of revenue from PBF	Annual revenues that the school receives through the PBF scheme	RBF audit
Brick walls	Number of brick walls in the school	RBF audit
Regular meeting with parents (d)	Dummy variable indicating whether regular meetings with parents are held	RBF audit
Mean overall test score	Mean test score of fifth grade students at the school	Own survey
Average score on French test	Mean French test score of fifth grade students at the school. Maximum score is 27.	Own survey
Average score on Math test	Mean math test score of fifth grade students at the school. Maximum score is 21.	Own survey
Shabunda district	Dummy variable indicating whether the school is located in Shabunda	RBF audit
RBF_TPerf	Index which measures teacher and team performance at baseline. It is the total sum of 35 indicator variables used as part of RBF to evaluate the quality of teacher and team performance. This list includes, for instance, an assessment whether the teacher uses report cards, uses notes to prepare for class, and tracks student attendance. A higher value indicates better teacher and team performance	RBF audit
RBF_QProg	Index which measures the quality and continuity of the teaching programmes. It is the total sum of 35 indicator variables used as part of RBF to evaluate the quality and continuity of the teaching programmes. This list includes, for instance, whether a file on each student exists and is passed on from one teacher to the next and whether shared teaching documentation exists	RBF audit
School climate	Perceptions of the overall social and learning climate of the school by headmaster and teacher, measured using nine Likert-scaled items with verbal labels for the answer categories. For example: “How clearly do your school leaders identify their goals for	Own survey

	teachers?" Answer categories are *not at all clearly, slightly clearly, somewhat clearly, quite clearly, and extremely clearly*. In the analysis, we use the average score of responses across the nine items. Source: Harvard Panorama Survey	
School leadership	Perceptions of the school leadership's effectiveness Nine Likert-scaled items, with verbal labels for the answer categories. For example: "How thoroughly do you feel that you know all the content you need to teach?" Answer categories are *not at all confident, slightly confident, somewhat confident, quite confident and extremely confident*. In the analysis, we use the average score of responses across the nine items. Source: Harvard Panorama Survey	Own survey
Teaching experience of fifth grade teacher	Number of years of teaching experience	Own survey
Teaching efficacy of fifth grade teacher	Faculty perceptions of their professional strengths and areas for growth, assessed using nine Likert-scaled items, with verbal labels for the answer categories. For example: "How thoroughly do you feel that you know all the content you need to teach?" Answer categories are *not at all confident, slightly confident, somewhat confident, quite confident and extremely confident*. In the analysis, we use the average score of responses across the nine items. Source: Harvard Panorama Survey	Own survey
Frequency feedback	Perceptions of the amount and quality of feedback faculty and staff receive, measured using five Likert-scaled items, with verbal labels for the answer categories. For example: "How much feedback do you receive on your teaching? Answer categories are *no feedback of all, a little bit of feedback, some feedback,* quite a bit of feedback, *a tremendous amount of feedback*. In the analysis, we use the average score of responses across the five items. Source: Harvard Panorama Teacher Survey	Own survey
Perceived school lead	Perceptions of faculty and staff relationships with school leaders. Ten Likert-scaled items. For example: "How friendly are your school leaders toward you?" Answer categories are *Not at all friendly, slightly friendly, somewhat friendly, quite friendly,* and *extremely friendly*. In the analysis, we use the average score of responses across the ten items. Source: Harvard Panorama Teacher Survey	Own survey
School climate	Perceptions of the overall social and learning climate of the school; measured using nine Likert-scaled items, with answer categories that have verbal labels. For instance: "On most days, how enthusiastic are the students about being at school?" Answer categories are *not at all enthusiastic, slightly enthusiastic,somewhat enthusiastic, quite enthusiastic, extremely enthusiastic*. In the analysis, we use the average score of responses across the nine items. Source: Harvard Panorama Teacher Survey	Own survey

Quality of school construction	On a scale from 1 to 3 assessed by enumerator. 1 corresponds to *temporary construction (bamboo, leaves, tarpaulin ...),* 2 indicates *semi-hard construction* (*boards, earth, banco, wood, sheets, ...)*, and 3 indicates *hard construction (cement, brick, stones, ...)*	Own survey

Table A.2: Explaining attrition with pupil, school, and teacher characteristics

	Probit
Pupil	
age (in years)	0.154*
	(0.089)
girl (d)	-0.902***
	(0.291)
mother is literate (d)	0.129
	(0.275)
father is literate (d)	-0.138
	(0.379)
receives help for homework (d)	0.007
	(0.161)
minutes spent on homework	0.006*
	(0.003)
Hours of labour	0.032*
	(0.019)
Experienced level of violence (1-5)	-0.025
	(0.054)
breakfast (frequency 1 to 5, most frequent)	0.045
	(0.069)
School	
Organised study time at school after school	-0.260*
	(0.155)
Teacher supervises during this time	-0.502**
	(0.197)
Ratio success at TENAFEP	0.264
	(0.226)
Ratio register at TENAFEP	-0.029
	(0.190)
Teachers per pupil	1.735
	(5.389)
Size of grade 5 class	0.001
	(0.002)
Number of vulnerable children	1.888
	(1.438)
Ratio girl/boy	0.038
	(0.221)
Ratio of revenue from RBF	-0.007
	(0.044)
Number of brick walls	0.000
	(0.017)
Regular meeting with parents (d)	0.169*
	(0.088)
Mean TENAFEP score	-0.034
	(0.023)

Shabunda district (d)	0.082
	(0.104)
RBF index on teacher and team performance	-0.005
	(0.006)
RBF index on quality and continuity of teaching	-0.017
	(0.011)
School leadership index	0.452***
	(0.151)
School climate index	0.117
	(0.189)
Teacher	
Grade 5's teacher years of experience	-0.002
	(0.004)
Average on Harvard teaching index	0.154
	(0.138)
Treatment group	-0.057
	(0.078)
N	2814

d) dummy variable. Standard errors in parentheses. P-value * < 0.1 ; ** < 0.05 ; *** < 0.01

Table A.3: Compliance to the intervention (in 2017)

	Semester 2	**Semester 3**
Books returned in good state	0.754	0.678
	(0.277)	(0.308)
Registers are used [core] (b)	0.667	0.622
	(0.477)	(0.490)
Stars system is used [core] (b)	0.663	0.63
	(0.301)	(0.311)
Pictures displayed in the classroom [core] (b)	0.8	0.778
	(0.405)	(0.420)
Proportion of book withdrawals	0.75	0.726
	(0.316)	(0.334)
Project documents are filed [core] (b)	0.778	0.8
	(0.420)	(0.405)
Stars obtained	0.681	0.629
	(0.284)	(0.310)
Weekly test (b)	0.64	0.53
	(0.283)	(0.257)
N	45	45

(b) binary variable. Standard deviations in parentheses.

Table A4: Heterogeneity in treatment effect by students' socio-economic status, gender, baseline score and age

	Dependent variable: Standardized test scores (end line)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
COVARIATES	Indigent		Female		Best half of the students		Older pupils	
	French	Math	French	Math	French	Math	French	Math
Treatment	0.368**	0.074	0.318*	0.053	0.439***	0.074	0.265*	0.033
	(0.152)	(0.091)	(0.168)	(0.102)	(0.164)	(0.115)	(0.147)	(0.098)
Covariate	0.192	0.189	0.041	0.180***	0.128	0.146*	0.062	0.135
	(0.159)	(0.138)	(0.083)	(0.067)	(0.110)	(0.087)	(0.104)	(0.099)
Interaction	-0.403*	-0.307	0.001	-0.034	-0.253*	-0.086	0.155	-0.001
	(0.209)	(0.189)	(0.110)	(0.091)	(0.152)	(0.118)	(0.116)	(0.105)
Treatment plus Interaction			**				**	
Observations	1498	1498	1498	1498	1498	1498	1498	1498
Adjusted r-squared	0.290	0.283	0.287	0.281	0.290	0.283	0.289	0.282

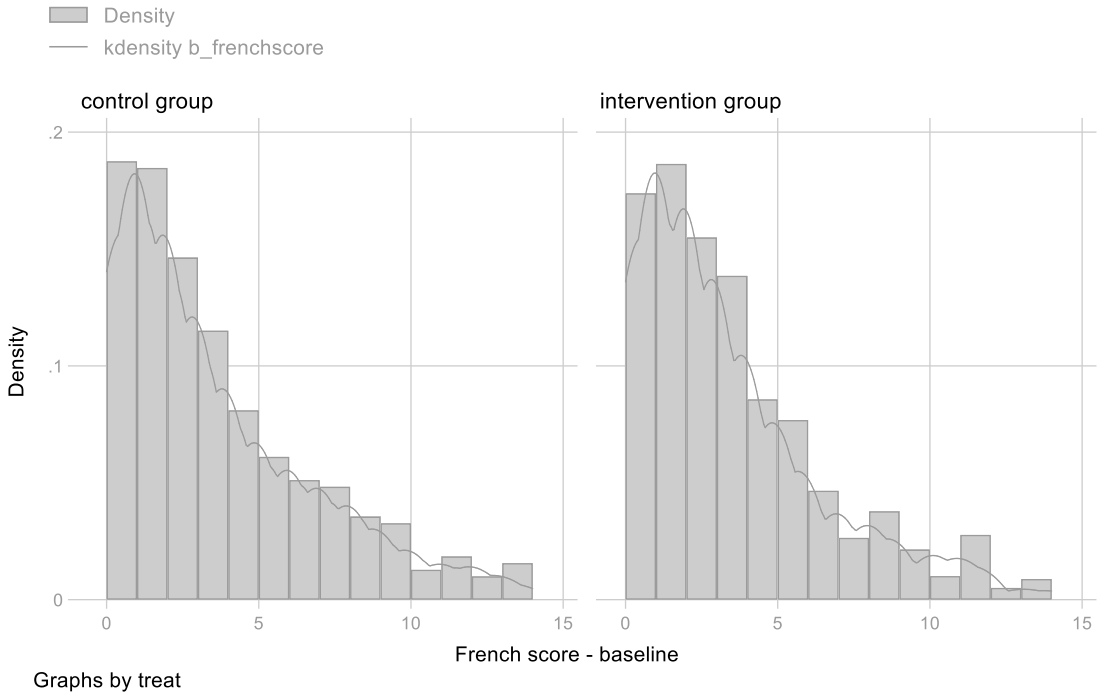
Note: Robust standard errors clustered by school in parentheses. All regressions include enumerator fixed effects, the four controls mentioned earlier and inverse probably weighting to account for attrition issues. p-value * < 0.1 ; ** < 0.05 ; *** < 0.01

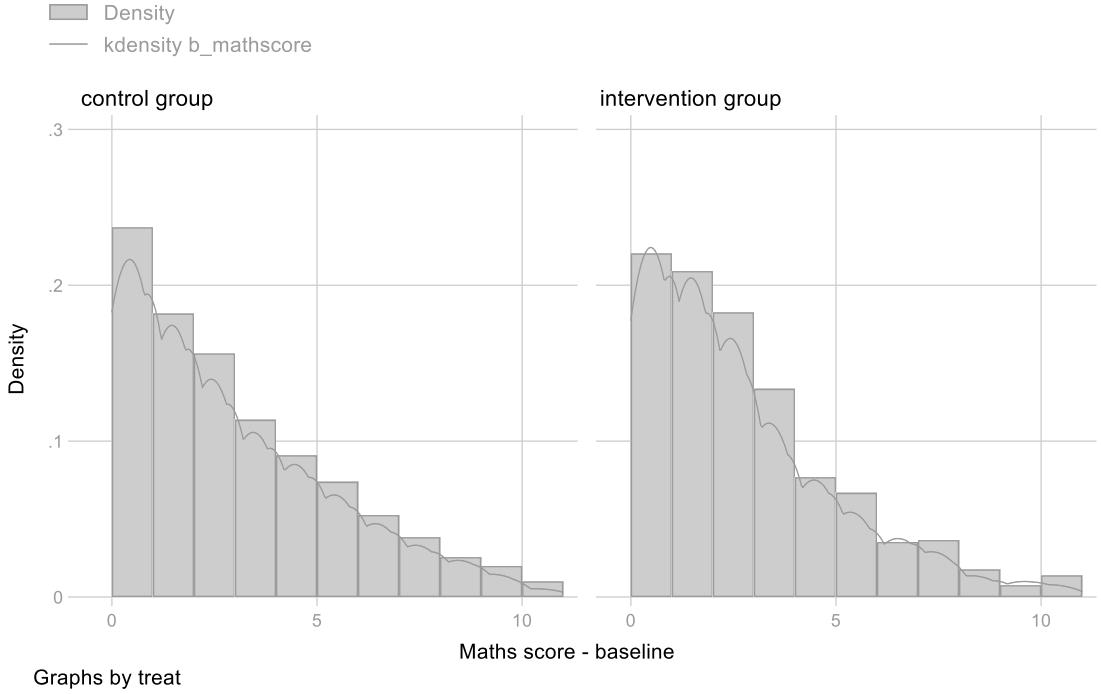
Table A5: Heterogeneity in treatment effect by teacher characteristics (all binary, with 1 for an above-the-median score)

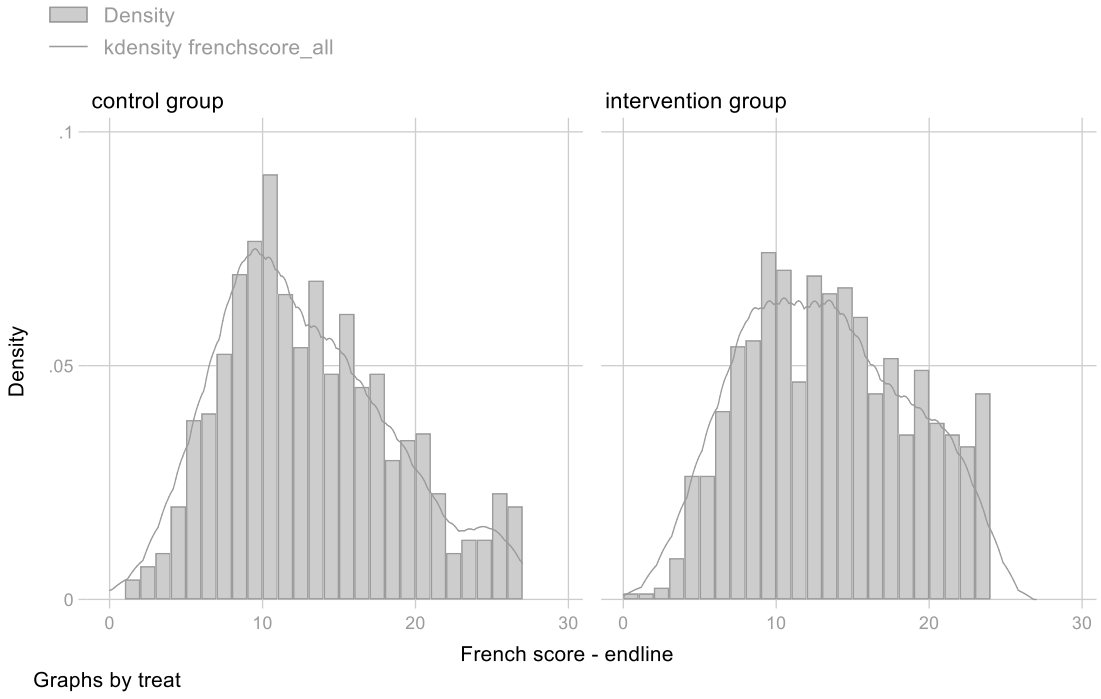
	Dependent variable: Standardized test scores (end line)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
COVARIATES	Teaching efficacy		Teaching experience		Teacher competence in French		Teacher competence in Math	
	French	Math	French	Math	French	Math	French	Math
Treatment	0.279	0.080	0.599**	-0.040	0.334	-0.024	0.314	0.202
	(0.218)	(0.138)	(0.267)	(0.159)	(0.234)	(0.171)	(0.273)	(0.179)
Covariate	-0.224	-0.117	0.163	0.102	0.166	-0.137	0.002	0.068
	(0.250)	(0.193)	(0.189)	(0.127)	(0.256)	(0.234)	(0.279)	(0.213)
Interaction	0.124	-0.057	-0.539*	0.141	-0.208	0.207	-0.135	-0.199
	(0.384)	(0.254)	(0.315)	(0.221)	(0.340)	(0.271)	(0.372)	(0.271)
Treatment plus Interaction								
Observations	1498	1498	1498	1498	1498	1498	1498	1498
Adjusted r-squared	0.290	0.283	0.287	0.281	0.290	0.283	0.289	0.282

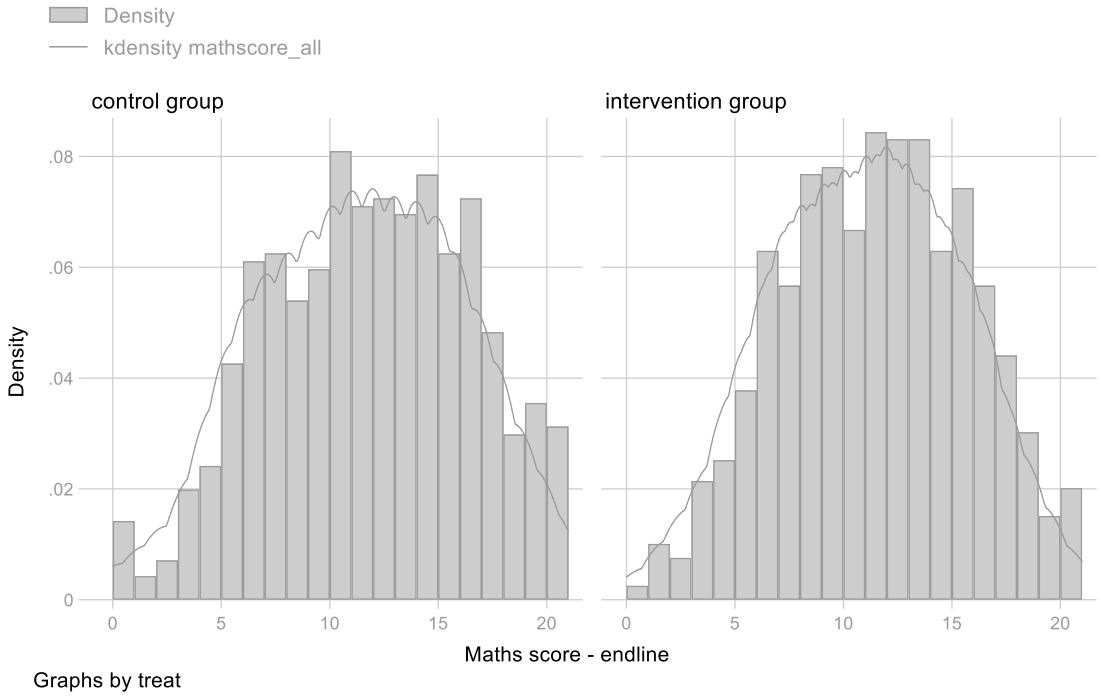
Note: Robust standard errors clustered by school in parentheses. All regressions include enumerator fixed effects, the four controls mentioned earlier and inverse probably weighting to account for attrition issues. p-value * < 0.1 ; ** < 0.05 ; *** < 0.01

Figures A1: distribution of test scores at baseline and endline






Density
kdensity frenchscore_all
control group
intervention group
Density
.1
.05
0
0
10
20
30
French score - endline
Graphs by treat


Density
kdensity mathscore_all
control group
intervention group
Density
.08
.06
.04
.02
0
0
5
10
15
20
Maths score - endline
Graphs by treat

Baseline test instrument

Cordaid BUILDING FLOURISHING COMMUNITIES

TEST DE PERFORMANCES DES ELEVES DE 5e ANNEE PRIMAIRE[19]

Noms de l'élève :……

Code de l'élève : ……

Nom de l'école : ……

Code de l'école : ……

Territoire : …… Date: …… **……**

Consigne/ Background :

Ce test dure 1 heure. Il n'est pas différent d'autres tests et examens que vous passez généralement en classe. D'ailleurs les questions sont adaptées au niveau de la 5e année primaire en RDC. Il est donc important de ne pas avoir peur, de lire attentivement et comprendre chaque question et d'y répondre. Les réponses sont proposées. A vous de trouver la bonne réponse qui convient. Ne cochez pas deux réponses pour une même question. N'oubliez pas d'écrire vos noms sur le papier.

Donner aux élèves l'occasion de poser toutes les questions possibles sur la façon de passer le test et de répondre aux questions posées. Ne pas lire ni expliquer les questions pour les élèves, seulement il importe de leur donner les consignes standards : choisir la vraie réponse parmi les alternatives proposées pour chaque questions après calcul et/ou réflexion. Chaque question a une seule réponse vraie parmi celles proposées)

Français

Série de questions 1:

Quel est le verbe qui est correctement conjugué pour remplacer les ……..dans les phrases suivantes :

1.1. Mon père et moi …… au marché

a) Va b) vont c) allons.

1.2. Le chauffeur …… à l'entrée du village

a) ralentit b) ralentis c) ralentissent.

[19] Ce test est tiré du MINESPSP/PASEC X, Programme d'Analyse des Systèmes Educatigs de la CONFEMEN – Opérations de fin d'année scolaire 2009-2010

Série de questions 2:

Lis les propositions suivantes et complète les ……. par un pronom personnel qui convient pour remplacer les mots soulignés:

2.1. Les élèves de 2e année pensent **<u>à leur examen</u>**, les élèves de 5e année …….. pensent.
2.2. **<u>Jean et moi</u>** allons au stade aujourd'hui ; …….. allons au stade aujourd'hui.

Série de questions 3:

Complétez les …. par c'est, s'est, ces, sait :

3.1. ……. fourmis ont du mal à transporter leur nourriture.
3.2. Maintenant ……..un plaisir pour moi de vous inviter.
3.3. Il ……..blessé au genou.

Dans une boîte de médicaments, on trouve la notice suivante : Lisez-là attentivement et répondez aux propositions.

PRIMALAN

INDICATIONS

4. Dérangements intestinaux et plus spécialement :
   - ✓ Diarrhées
   - ✓ Vomissements.

5. *POSOLOGIE :*
   - ✓ Adultes : 1 à 6 comprimés par jour
   - ✓ Enfants de 3 à 5 ans : ½ comprimé deux fois par jour.
   - ✓ Enfants au-dessus de 5 ans : ½ comprimé 2 à 4 fois par jour.

   Ce médicament est à prendre au début du repas en avalant avec peu d'eau, sans croquer.

Série de questions 4 :

Propositions :

4.1. Est- ce que ce médicament guérit la diarrhée ? (cochez la bonne réponse)
a) Oui ☐
b) Non ☐
c) Le texte ne le dit pas ☐

4.2. Un adulte peut prendre 8 comprimés par jour (cochez la bonne réponse)
a) Oui ☐
b) Non ☐
c) Le texte ne le dit pas ☐

4.2. Un enfant qui a plus de 5 ans peut prendre (cochez la bonne réponse)
a) 1 à 6 comprimés par jour ☐
b) 1 comprimé deux fois par jour ☐
c) ½ comprimé deux à quatre fois par jour ☐
d) Le texte ne le dit pas ☐

Série de questions 5 :

Lis le texte et complète-le à l'aide des mots suivants (attention, il y a un mot de trop).

à – berger – vache – avouer – au – rentré – fusils – une fausse alerte.

5.1. Samedi, vers 9h du matin, trois petits garçons ont mis des masques pour ressembler à des lions. Ils sont allés ……bord du marigot et se sont mis à rugir lorsqu'un …… est arrivé avec son troupeau. Affolé, celui-ci est ……au village et a informé les éleveurs. Ceux-ci ont fait appel ……la gendarmerie, aux gardes, aux militaires.

5.2. Tout le monde est arrivé armé de ……. Heureusement les trois garçons avaient enlevé leurs masques et pêchaient tranquillement. Interrogés, ils ont été obligés d'…… leur blague. Les gendarmes les ont grondés et ont confisqué leurs masques.

Mathématiques

Série de questions 6 :

Range les nombres du PLUS PETIT au PLUS GRAND. ***Ecris ton rangement sur les points***

6.1) 8 210 8 120 8 217 8 107

……..……
……..…

6.2) 14 940 1 449 14 409 10 049

……
……

Range les nombres du PLUS GRAND au PLUS PETIT. ***Ecris ton rangement sur les points***

6.3) 35,7 25,9 35,8 35,6

……
……

6.4) 0,82 0,15 0,25 0,90

……
……

6.5) 794,9 795,5 794 79,4

……
……

Série de questions 7:

Effectuez les opérations suivantes et entoure la bonne réponse :

7.1) 54 + 20,7= ……

747
74,7
61

7.2) 46,3 + 23 + 178,25 =……

183,11
871,25
247,55

7.3) 134 – 82,48 = …….

52,48
51,52
81,14

Série de questions 8:

Entoure la plus petite fraction de chaque série

8.1) $\frac{8}{4}$ $\frac{8}{32}$ $\frac{8}{16}$ $\frac{8}{2}$

8.2) $\frac{21}{7}$ $\frac{3}{7}$ $\frac{9}{7}$ $\frac{14}{7}$

Série de questions 9:

Estime la quantité d'eau dans les récipients dessinés ci-dessous (entoure la bonne réponse) :

9.1) La bouteille pleine fait un litre. Sur le dessin, la quantité d'eau est donc de :…….

a) 1 litre
b) 0,75 litre
c) 0,25 litre

9.2) Le fût plein fait 10 litres. Sur le dessin, la quantité d'eau est donc de :……..

a) 3 litres

b) 1 litre
c) 5 litres