

Entrepreneurship Education and Teacher Training in Rwanda*

Moussa P Blimpo
World Bank

Todd Pugatch
Oregon State University

March 2020

Abstract: We assess, via an experiment across 207 secondary schools, how a comprehensive teacher training program affects the delivery of a major entrepreneurship curriculum reform in Rwanda. The reform introduced interactive pedagogy and a focus on business skills in the country's required upper secondary entrepreneurship course. In addition to the government's standard training, a random sample of schools received intensive training organized by an NGO for two years. The training consisted of (i) six training sessions during school breaks, (ii) exchange visits each term where teachers provided feedback to their peers, and (iii) outreach and support from NGO staff at least twice per year. The program increased teachers' use of active instruction, consistent with the reform's features. These effects on pedagogy did not translate into improvements in student academic outcomes or skills. Treated students increased their participation in businesses by 5 percentage points, or 17% of the control mean, with a commensurate decrease in wage employment, and no effect on overall income. These results suggest substitution between entrepreneurship and employment among students in treated schools.

Keywords: entrepreneurship education; teacher training; secondary school; pedagogy; randomized control trials; Rwanda

JEL Codes: I25; I26; I28; J24; O12; O15.

* We are grateful to all the partners involved in this project. Particular thanks are due to Marius Chabi, Kris Cox, Doug Kirke-Smith, Katie MacAulay, Emmanuel Musafiri, Jean Bosco Nkurikiye, Jean Aime Nsabimana, and Phillip Okull of Innovations for Poverty Action; Maggie Appleton, Rogers Kamugisha, Meghan Mahoney, and Laura Smith of Educate!; Antoine Mtsinzi and Claudien Nzitabakuze of the Rwanda Education Board; and Jacques Sezikeye of Akazi Kanoze Access. Henry Aviomoh, Roberta Nilson, and Ruifan Shi provided excellent research assistance. The project has been enrolled in the AEA Trial Registry (AEARCTR-0001030, Pugatch and Blimpo 2016). This project received human subjects research approval from the Innovations for Poverty Action Institutional Review Board 0004745 USA, protocol #13918. This work was supported by the Jameel Poverty Action Lab (J-PAL) Post-Primary Education (PPE) Initiative; the Innovation Growth Lab (grant GR-5209); and the IZA/DFID Growth and Labor Markets in Low-Income Countries (GLM-LIC) Phase IV (grant C4-RA3-416). The findings, interpretations, and conclusions expressed in this report are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. All errors are due only to the authors.

Table of Contents

I. Introduction	3
II. Context and Program Description	7
II.1. Context.....	7
II.2. Intervention and Implementation	9
II.3. Theory of Change	10
III. Research Design	10
III.1. Sampling and Randomization.....	11
III.2. Data	12
III.3. Summary Statistics and Identification.....	14
III.4. Empirical Strategy	15
IV. Results	17
IV.1. Main findings.....	17
IV.1.1. Uptake and implementation: do schools adhere to the in-service training?.....	17
IV.1.2. Are teachers adjusting their pedagogy?.....	19
IV.1.3. Has the intervention improved student learning outcomes?.....	20
IV.1.4. Do we observe more entrepreneurial activities and higher earnings?	21
IV.2. Heterogeneity and Mechanisms	23
IV.2.1. Heterogeneous treatment effects	23
IV.2.2. Mechanisms	25
V. Conclusion	26
References	28
Appendices.....	31
Appendix 1: Entrepreneurship Syllabus	31
Tables	34
Figures.....	62

I. Introduction

With the fastest growing and youngest population among developing regions, Africa needs to create millions of jobs annually to keep pace with labor force growth, amidst uncertainty about the future of work (World Bank 2018b). Education systems in African countries have produced far more job seekers, often with inadequate skills, than entrepreneurs who create opportunities for others (Awusu and Blimpo 2019). Youth account for 60 percent of Africa's unemployed. In Rwanda, 72 percent of employed youth work for family firms or are self-employed, often in low productivity activities in the informal sector (Filmer and Fox 2014; African Economic Outlook 2016; Fox, Senbet, and Simbanegavi 2016). Increasing skills and productivity among the self-employed is thus viewed as a policy priority to address youth employment challenges in the region.

While the poor quality of education has garnered much attention in recent years (World Bank 2018a), more attention is needed to better understand the link, or lack thereof, between schooling, skill acquisition, and labor market outcomes. A specific reason is inadequate curricula, which fail to provide students in secondary and tertiary schools with the necessary entrepreneurial skills. In recognition of the challenging youth labor market, in 2016 Rwanda reformed its primary and secondary curricula, including the required secondary entrepreneurship course, by introducing interactive and student-centered pedagogy with a focus on practical skills. Successful reforms require adequate implementation, however. A survey of the literature found that “implementing student-centered instruction effectively requires skills well beyond those of a great many teachers in developing countries” (Murnane and Ganimian 2014, p. 42).

This study focuses on this potential implementation gap by examining how a comprehensive teacher training affects the delivery of the reform. The training, organized by an NGO, was subject-specific (entrepreneurship), consisted of six multi-day training sessions spread evenly over two years, incorporated peer feedback meetings each term, and included follow-up support by NGO staff.

We evaluate an experiment in which 103 schools were randomly selected to receive the training for two years. The program covered more than 260 teachers and 6,800 students. A control group of 104 schools received the curriculum and the default government training. The default training lacked each of the features of the intervention delivered in the treatment schools.

We therefore test whether such comprehensive training can improve delivery of a newly adopted, active learning curriculum. How effective is comprehensive teacher training in changing teacher pedagogy, building student entrepreneurial skills, and promoting student economic activity? Our analysis follows the Stage 1 Registered Report accepted by this journal (Blimpo and Pugatch 2019).

Two years after the program began, we report program impacts on teacher pedagogy, student learning, and student economic activities. First, on average, treatment schools sent teachers to more than five out of the six scheduled trainings (88%) and four out of every five exchange visits (81%). Treatment increased scheduled Skills Labs (double periods modeled on science laboratory classes) by 52 percentage points, a more than six-fold increase relative to the control group. This high uptake translated into measurable changes in teacher pedagogy. The intervention increased active instructional time, when expected, by 10 percentage points, or 20% of the control mean. The effect was negative when restricted to Skills Labs, double class periods in which interactive pedagogy should feature most prominently. However, this result may be explained by self-selection into Skills Labs among highly motivated control teachers. Overall, active instructional techniques such as group discussion, research, case studies, debates, and role playing increased by 6 percentage points, or 19% relative to the control group.

Second, despite high uptake and changes in pedagogy, we find no effect on students' learning outcomes measured by scores on high-stakes exams, including the national entrepreneurship exam. We also find no effect on our own test of financial literacy and entrepreneurship knowledge. Similarly, we find no effect on students' non-cognitive skills, including educational and professional aspirations, grit, and locus of control.

Third, the intervention increased student participation in revenue-generating businesses, but reduced wage employment by a commensurate amount. Business participation increased by 5 percentage points, a 17% increase over the control group mean of 30%. The effect was driven by student business clubs, highlighting schools as the channel of the increase. Student employment decreased by 5 percentage points, leaving overall income unchanged. These results suggest substitution between entrepreneurship and wage employment among treated students. Although business activities increased, we observed a decrease in business profits conditional on participation.

Fourth, we assess whether the effect of the intervention varied for different subgroups of teachers and students. We find no differential effect on business involvement, employment, or income by student or teacher gender. Higher baseline exam scores associated with higher treatment effects on student employment. Students from above-median socio-economic status households earned less income in response to treatment. We find no differential effect by teacher experience, while teacher qualification is associated with a differential negative effect on business participation.

Finally, we explore potential mechanisms by interacting the treatment with individual components of the interventions (training and exchange visits), and with teaching practices. We find no evidence that any specific component of the program or pedagogy in isolation affects business participation, employment, or income.

Taken together, the findings constitute promising signs of pedagogical change in response to the program. The results may also indicate that pedagogical changes bear fruit slowly, despite an intensive, well-crafted intervention with high take-up over two years. Teachers' adherence to the program was remarkably high considering that trainings were scheduled during school breaks and not formally mandated. Teachers clearly have demand for training and support when implementing an ambitious new curriculum. The observed changes in pedagogy demonstrate teachers' willingness to try new practices in an attempt to improve their students' outcomes.

Increased student engagement in business activities is also an early sign that when implemented as intended, the curriculum can be successful in promoting entrepreneurship. Outcomes were measured before students had completed secondary school, an important limitation to consider when interpreting results for student economic activity. However, our failure to find robust increases in academic performance or student skills raises concerns that the intervention successful built a foundation for longer-term change.

This study makes four main contributions to existing knowledge. First, we add to the literature on teacher training in developing countries by providing evidence from secondary school teachers. In-service teacher training, the approach we evaluate here, has shown promise to improve teaching of traditional curricula (Angrist and Lavy 2001; Paul Glewwe et al. 2013; Cilliers et al. 2019). The details of such programs matter, however. The program we study focuses on a single subject, incorporates lesson enactment, and includes follow-up visits. Each of

these elements is associated with positive student outcomes (Popova et al. 2018; World Bank 2018a), but is lacking in the training received by the control group. However, prior knowledge about teacher training in developing countries stems largely from primary schools (Null et al. 2017).

Second, our work adds to emerging evidence on how to improve teaching quality by altering pedagogy. Improving pedagogy has been identified as a leading mechanism for the success of education interventions in developing countries (Evans and Popova 2016). Rwanda's curriculum reform represents a major shift in pedagogy from traditional knowledge acquisition to student-centered, active learning. Many other efforts to alter pedagogy have a similar goal of promoting active learning, such as the early grade literacy program studied in Kerwin and Thornton (2015) and the teacher coaching program studied in Bruns, Costa, and Cunha (2018). Nonetheless, not all such efforts have been successful. In an experiment promoting active learning in secondary school mathematics, control group students learned more than the treatment group, despite 40 hours of training for treated teachers (Berlinski and Busso 2017). The circumstances under which pedagogical changes improve student outcomes, therefore, remain an area open for inquiry.

Third, this study sheds more light on how governments can best implement curricular reform. Teachers may be unable to change curricula effectively without additional training and support. Indeed, education interventions often depend crucially on such complementary inputs for success (P. Glewwe and Muralidharan 2016; Mbiti et al. 2019). This study can inform curricular reform efforts across many contexts. Within Rwanda, results are directly relevant for potential scale-up as entrepreneurship is a required subject and government employees delivered the training by the end of the program. Similar reforms are occurring or under discussion in several other African countries, including Ethiopia, Kenya, Mauritius, and Zambia.

Finally, we add to the thin evidence on school-based entrepreneurship education. To our knowledge, the only experiment of school-based entrepreneurship training in sub-Saharan Africa was conducted in Uganda by Educate!, the international NGO partnering with Rwanda's Ministry of Education on this program (Educate! 2014). In fact, we know of only two other experiments worldwide of school-based entrepreneurship education, and neither focused on secondary students (Premand et al. 2012 and Alaref, Brodmann, and Premand 2019 on university students in Tunisia; and Huber, Sloof, and Van Praag 2014 on primary students in the

Netherlands). Other experiments to encourage entrepreneurship, such as Blattman, Fiala, and Martinez (2014, 2018) or Alibhai, Buehren, and Papineni (2016), target a mostly older population that has already left school.¹ Our focus on secondary students is promising because early skill acquisition could lead to high returns.²

Even if the program proves successful at improving outcomes and is cost-effective, scale-up may present additional challenges. Despite public delivery of the bulk of the training program, the intervention included additional elements (peer feedback meetings, referred to as “exchange visits,” and outreach) led by a well-managed NGO. Other studies have found education interventions to be less effective when implemented by the government than by NGOs (Kerwin and Thornton 2015; Bold et al. 2018). Indeed, some large-scale, publicly managed teacher training programs fail to have positive effects (Loyalka et al. 2019). Readers should bear these caveats in mind when interpreting results.

II. Context and Program Description

II.1. Context

The education system in Rwanda: Rwanda’s education system consists of 6 years of primary school (grades P1-P6), 3 years of lower secondary (S1-S3), 3 years of upper secondary (S4-S6), and various tertiary options. The academic year runs from January through November, split into three terms. The primary grades are compulsory. All Rwandan secondary students are required to enroll in the entrepreneurship course throughout grades S1-S6.³ The requirement has been in place since 2009, making Rwanda the “site of one of the most extensive efforts to promote youth entrepreneurship in the world” (Honeyman 2016, p. xii).

Gross enrollment in Rwandan secondary schools is 42 percent for girls and 39 percent for boys (World Bank 2015). Of Rwanda’s 1,543 secondary schools, 30 percent are public, 40

¹ Vocational training can also encourage youth entrepreneurship, though programs often focus on trade skills, not business creation (e.g., Card et al. 2011; Hicks et al. 2015).

² Several countries in Sub-Saharan Africa, such as Ghana, Kenya, Mozambique, and South Africa, also offer early-age entrepreneurship education (Robb, Valerio, and Parton 2014; Bux 2016).

³ Rwanda also offers Technical Vocational Education and Training (TVET) in a separate system of secondary schools. TVET focuses on occupation-specific training in fields such as office management, accounting, and agriculture. TVET students are not subject to the entrepreneurship requirement. By contrast, students in the required entrepreneurship course are enrolled in “general” secondary schools with an academic focus. TVET schools account for 16 percent of Rwanda’s secondary school enrollment (Rwanda Ministry of Education 2016, p. 14).

percent are Catholic, with the remainder run by other religious or private institutions (Rwanda Ministry of Education 2016, p. 38). The completion rate for upper secondary school in Rwanda was 18 percent in 2015 (UNESCO Institute for Statistics 2018).

Table A1 presents data on youth schooling and economic activity from the 2012 Rwandan Census. Nationally, 63 percent of youth aged 15-19 are enrolled in school. Among youth aged 20-24, the enrollment rate falls to 24 percent, indicating that many youth transitions from secondary school to the labor market at these ages. University attendance is about 4 percent among 20-24-year olds, indicating that relatively few students continue their studies after secondary school. Among 15-19 year olds, 25 percent are employed, with the employment rate among 20-24 year olds rising to 54 percent.⁴ Again, this indicates that these age ranges mark the transition from school to the labor market for many youth. Among the employed, most are self-employed or work for a family firm (74 percent). Wage labor is therefore scarce for Rwandan youth, emphasizing the importance of entrepreneurial knowledge and skills for economic success. Most (67 percent) of the employed are also in the agricultural sector.

The revised entrepreneurship curriculum: This project focuses on training teachers to deliver the entrepreneurship course for upper secondary grades S4-S6 (10th-12th grade). In 2016, the government reformed the primary and secondary curricula to focus on building skills through active learning. They called the reformed courses the “competency-based curriculum,” a student-centered approach, contrasting them with the previous “knowledge-based curriculum” that used traditional teaching practices. The entrepreneurship course required of all secondary school students was part of this reform. Mastery of entrepreneurship and other required subjects, as measured by formal exams, is required to complete secondary school.

The government reformed the upper secondary entrepreneurship curriculum with consultation from Educate!. The reformed course covered the full cycle of business creation and development, including product development, registration, and legal issues, marketing, accounting, and customer relations. The course ranged from covering specific topics and skills (e.g., a lesson on “marketing materials”) to more general skills (e.g., “effective communication,” “setting goals”). We list the expected key competencies and Skills Lab topics for each grade in Appendix 1. In addition to promoting greater interaction between teachers and students, the

⁴ Employed refers to those who answered “Yes” when asked “Aside from your own housework, did you work at least 1 hour during the last 7 days preceding the census night?”

reformed course included detailed plans for weekly “Skills Labs,” based on the laboratory science model, in which students practiced business skills through role-play and group projects. Class periods for Skills Labs expanded from 40 to 80 minutes to accommodate the new format.

The new course also encouraged students to form “student business clubs” to start and run school-based businesses. The purpose of these extracurricular clubs was to allow students to practice their entrepreneurial skills in revenue-generating firms of their own creation.

II.2. Intervention and Implementation

The requirement that all secondary school students enroll in the revised entrepreneurship course prevents a direct evaluation of the entrepreneurship curriculum, as there would be no comparable group of students unexposed to the new curriculum. Instead, this project focuses on the extent to which comprehensive teacher training can improve curricular implementation. The intervention tested in this project consisted of the following components:

In-service teacher training: entrepreneurship teachers received multi-day training sessions each academic term beginning April 2016 through January 2018. Each of the six sessions was held during holidays between terms and lasted four days. The training covered pedagogical strategies for implementing the revised entrepreneurship curriculum. The training emphasized lesson planning, engaging students in classroom discussions, encouraging students to create entrepreneurship “portfolios” of their work, and assisting student business clubs to form and grow. The training culminated in a “mock day” in which teachers rehearsed upcoming lessons. Government trainers led the trainings. These trainers were trained by Educate! in a “train the trainers,” or “cascade,” model.⁵

Exchange visits: teachers participating in the intervention visited each other’s schools to learn from and provide feedback to their peers. Each term, beginning in June 2016 through March 2018, teachers and a district education official observed a colleague teaching an entrepreneurship lesson. After the lesson, teachers conducted a roundtable discussion to share

⁵ Government trainers received an initial 5-day intensive training from Educate!, with “refresher” trainings each term. All trainers who participated in the cascade model had previously received training from government before receiving training from Educate!. It is therefore possible that trainers overlapped between treatment and control schools, though we lack data to verify. Nonetheless, the trainings for control schools were not specific to entrepreneurship, leaving less opportunity to introduce techniques emphasized in treated schools into the control group trainings. For instance, Skills Labs were unique to the entrepreneurship course, and would not have been relevant to the non-entrepreneurship teachers attending control group trainings.

their observations and discuss pedagogical strategies. Teachers met in groups of 2-3, with the host school rotating each term.

Outreach and support: teachers received ongoing outreach to support their implementation of the curriculum. Youth Leaders, hired and trained by Educate!, visited schools participating in the intervention at least twice per term. The visits included product-making demonstrations (e.g., for household goods such as soap or candles) co-taught with the teacher, advising student business clubs, classroom observation, participating in teacher exchange visits, and addressing any other concerns. Student business clubs were encouraged to submit their ideas to regular business competitions held for treated schools.

The study focused on the cohort entering S4 (10th grade) in 2016, with training provided to this cohort's entrepreneurship teacher as they progressed to S6 (12th grade).⁶ The control group received the new entrepreneurship curriculum with only the standard government training on the competency-based curriculum, which was not specific to entrepreneurship. Teachers in control schools did not receive the intensive training, exchange visits, or outreach provided to treatment schools.⁷

II.3. Theory of Change

Despite the multiple elements of the program, the theory of change underlying the intervention is simple. Participation in training and support activities increases teacher adherence to the new curriculum and alters classroom pedagogy. As a result, students acquire an associated set of skills. They apply these skills in entrepreneurial and other economic activities. Figure 1 presents this theory of change as a series of numbered hypotheses. The empirical analysis revolves around testing each of these four set of hypotheses.

III. Research Design

The research compares two approaches to delivering a newly adopted, active learning curriculum. While many RCTs compare a new curriculum bundled with teacher training against

⁶ Up to two entrepreneurship teachers from each treated school were invited to each training.

⁷ Government training was scheduled for 10 days in 2016, with refresher sessions to be held in subsequent years. Each district could set the details of these sessions. We do not have data on implementation but suspect that training quantity and quality varied across districts, based on uneven responses to queries with district officials. Even if implemented as intended, training in control schools differed from treatment by occurring for fewer days, without NGO training and input, and without a standardized training curriculum.

a counterfactual that has neither (Banerjee et al. 2007; Lucas et al. 2014; Berlinski and Busso 2017), this study’s design holds the new curriculum constant and varies teacher training and support to implement it. The default government training in the control group represents the public status quo, making the comparison with treatment highly relevant for policy.⁸ The outreach component builds on studies of other training programs with similar follow-up for trained teachers (e.g., Beuermann et al. 2013; Abeberese, Kumler, and Linden 2014; Piper and Zuilkowski 2015). The exchange visits resemble a form of teacher coaching (e.g., Bruns, Costa, and Cunha 2018; Albornoz et al. 2018; Cilliers et al. 2019), except that trained teachers received feedback from peers rather than professional coaches.

III.1. Sampling and Randomization

The research design is a cluster randomized control trial, with treatment assigned at the school level.⁹ Our sample frame included 211 schools, spread across 11 districts in 3 of Rwanda’s 5 provinces. We randomly assigned 106 schools to treatment and 105 to control, stratifying treatment by district and public/non-public status of the school (i.e., across 22 strata).¹⁰ Randomization was conducted privately by the researchers, without any re-randomization.

Four schools refused to participate in the study, leaving 103 treatment and 104 control schools. Additionally, a miscommunication between the research team and project implementers led to one control school receiving the intervention, while two treatment schools did not receive the intervention. Although these discrepancies affect only 3 of 207 schools, we use initial random assignment in all analyses. Accordingly, all estimates should be interpreted as the intention to treat (ITT). Figure 2 maps the study design.

⁸ Another policy-relevant question would be to assess the effectiveness of the public status quo. While the rollout of the Rwandan curriculum reform does not allow for such evaluation, the quality of training and support are often key to the success of large scale reforms. For example, Blimpo et al. (2015) experimentally compared the roll-out of a new school governance program with and without training in The Gambia. The rollout without training and support had no impact, while the intervention with training and support reduced teacher and student absenteeism.

⁹ Lower levels of randomization, such as the classroom, are infeasible for two reasons. First, there was a single entrepreneurship teacher for the entering S4 cohort in 71 percent of schools, making school and classroom randomization equivalent for the bulk of the sample. Second, even if classroom randomization were feasible in a larger proportion of schools, the likelihood of spillovers across teachers and classrooms within the same school would contaminate treatment. Higher levels of randomization, such as the district, would also create problems because with 11 study districts, the small number of clusters would risk confounding treatment with district-level shocks. The school level therefore balances the tradeoff between the risk of treatment contamination and the need for a large number of clusters.

¹⁰ About half the schools (102) are non-public (private, government-aided, or religious). Although this is an interesting dimension to explore, we exclude it from our analysis plan to limit the number of outcomes considered in our analysis of treatment effect heterogeneity.

Blinding participants to treatment status was not possible, due to the nature of the program. For instance, teachers knew they had been invited to training sessions or exchange visits. However, program staff were instructed not to volunteer details of research design to participants, such as the division of the sample into treatment and control schools, or the study hypotheses. The sample size provided 80 percent statistical power to identify effects of 0.15-0.26 standard deviations in test scores and 7-12 percentage points in business ownership, depending on the assumed intraclass correlation (ICC).¹¹ These effect sizes fall within the range of positive outcomes found in many studies of education interventions.

III.2. Data

We use data collected from all schools, including survey data from teachers, students, and classroom observations. Data collection procedures were identical for treated and control schools. We also use administrative data obtained from the government and (for treated schools only) from Educate!. This project did not have a pilot.¹²

¹¹ All power calculations preceded endline data analysis. We assumed a test size of 5 percent and 80 percent power. We set sample sizes of 105 schools and 15 students per school, consistent with the research design. We used ICCs of 0.09, 0.25, and 0.43, corresponding to the observed baseline ICCs of student employment during the school holiday, business ownership, and S3 exam scores, respectively. For continuous outcomes, we assumed normalizing to the control group mean and standard deviation. For outcomes measured in proportions, we assumed a baseline proportion of 0.25, roughly in line with the share of students owning a business (0.22) or having a job last school holiday (0.27) in the baseline survey. Under these assumptions and ICCs, the sample size is sufficient to detect effects of 0.15, 0.21, and 0.26 standard deviations for continuous outcomes, and effects of 7, 10, and 12 percentage points for outcomes in proportions. These effect sizes are arguably conservative, as they do not account for increased precision from including baseline outcomes.

¹² We also conducted two midline surveys on subsamples of schools, in October 2016 and June 2017. The first midline included a subsample of 82 schools (38 control, 44 treatment). The imbalance between treatment and control was due to economizing on travel costs to visit schools within proximity of each other. The second midline survey was in a subsample of 80 schools (60 treatment, 20 control). The oversampling was deliberate in order to include all treatment schools from the first midline; all other schools differed between midline surveys. Each midline included a teacher survey, a survey of a subsample of students surveyed at baseline, and a classroom observation. We produced reports for government and the implementing NGO for each midline survey. The design and results of the midline played little role in the endline design, as most items in the endline questionnaire also appeared in the baseline (see bottom of Tables P3-P7 for mappings between baseline and endline variables). However, the midline surveys did influence the endline classroom observation. Based on the first midline survey, we expanded the duration of the observation (from 40 to 52 minutes) and tweaked some category definitions in the Stallings instrument (for instance, “Practice and Drill” became “Repetition of Facts From Memory”). We also decided to split the analysis by halves of the observations. Because the sample size and design make the midline data less appropriate for formal analysis, we do not include it in this submission or in the pre-analysis plan registered at the AEA Trial Registry. However, we may include results from midline surveys in the Stage 2 submission, labeled clearly as exploratory analyses.

The baseline survey occurred at the beginning of the 2016 academic year, before the intervention began. We surveyed the head teacher, the S4 entrepreneurship teacher (one was chosen randomly when a school had multiple S4 entrepreneurship teachers), and 15 randomly selected S4 students.¹³ Surveys covered school characteristics, perceptions of effective teaching practices, demographics, student labor market and economic activity, entrepreneurship knowledge, and non-cognitive skills.

The endline survey occurred from July-September 2018, during the entering cohort's final year of secondary school (S6). The endline included student, teacher, and head teacher surveys, and a classroom observation. We conducted the endline while students were still enrolled in order to maximize our ability to find students from the baseline, although it prevents us from observing any post-secondary outcomes. We surveyed all students from the baseline, including extensively tracking out-of-school students.¹⁴ The teacher survey included the S6 entrepreneurship teacher and the teacher surveyed at baseline¹⁵ and the S6 entrepreneurship teacher when this teacher differed from the baseline. We also observed the entrepreneurship class of each surveyed teacher. If possible, we observed a Skills Lab.¹⁶

Figure 3 presents the timeline of the project and research. The Kigali office of Innovations for Poverty Action (IPA) conducted all surveys. The baseline survey was conducted using paper records, with daily audits by field managers to ensure proper completion of surveys. Endline surveys used tablets for data entry in the field. Digitized data were checked for consistency by a Senior Research Associate in the IPA office. We also have administrative data on teacher training attendance to measure take-up and administrative data on student exam performance and school completion.

For the analysis, we do not impute missing values for any dependent variables. For covariates such as baseline outcomes, we replace missing values with the control group mean

¹³ Students were selected randomly from rosters submitted prior to baseline visits. Some schools had fewer than 15 students enrolled in S4, leading to a sample smaller than the expected 3,105.

¹⁴ We also sampled additional students not surveyed at baseline in order to sample 15 students during each school visit. We will not use data on these additional students in the analysis because they lack baseline outcome data and because our tracking of out-of-school students was so successful.

¹⁵ In cases where this was the same teacher, we surveyed one teacher per school. In cases where these were different teachers, we attempted to survey both, although time constraints sometimes prevented completion of two surveys.

¹⁶ Not all classroom observations were Skills Labs due to teacher noncompliance and scheduling difficulties with school visits. For this reason, we test for differences in classroom behavior overall and in Skills Labs. We also analyze the scheduling of Skills Labs as a separate outcome.

and include a dummy for missing in the regression (Haushofer and Shapiro 2015). To deal with outliers, we winsorize all financial variables (e.g., income, savings) at the 99th percentile.

III.3. Summary Statistics and Identification

Table 1, Panel A presents sample sizes. The sample contains roughly equal numbers of treated and control schools, teachers, and students. Teachers added to the sample at endline ensure that we have data on the study cohort’s current entrepreneurship teacher, in cases where this teacher is not the same as baseline. Attrition from the endline survey was minimal for students (7%) and somewhat large for teachers (over 30%).¹⁷

To assess whether attrition varies by treatment status, we estimate the following equation:

$$A_{isg} = \alpha + \beta T_{sg} + \gamma_g + \varepsilon_{isg} \quad (1)$$

where i indexes students; s indexes schools; and g indexes strata. The strata are district-school type cells, where school types are public and non-public. In this equation, A is an indicator for being absent in the endline; T is an indicator for assignment to treatment; γ is a stratum fixed effect; and ε is an error term. Because randomization occurred within strata, the strata fixed effects ensure that treatment assignment T is unrelated to the error term. The coefficient β measures the difference in attrition according to treatment assignment. We find that attrition is orthogonal to treatment assignment (Table 1, Panel B). Nevertheless, we construct and report Lee (2009) bounds for all our main results.

We next analyze our key identifying assumption that treatment and control groups are balanced at baseline. To do so, we compare means of baseline variables, with the variables chosen in accordance with a pre-analysis plan submitted to the AEA Trial Registry prior to analyzing the data. For each variable, we present unadjusted means and standard errors by treatment status. To formally test for differences, we estimate the following equation:

$$X_{oisg} = \alpha + \beta T_{sg} + \gamma_g + \varepsilon_{isg} \quad (2)$$

¹⁷ Some teacher attrition was due to survey budget constraints. In cases where the current entrepreneurship teacher of the study cohort differed from the baseline teacher, there was sometimes insufficient time to survey both.

where X_0 is a baseline characteristic of students, teachers, or schools; and all else is as in equation (1). The associated p -value will be our test for equality of means.

We also look for systematic balance by regressing the treatment indicator on multiple baseline variables:

$$T_{sg} = \alpha + X_{0isg}\beta + \gamma_g + \varepsilon_{isg} \quad (3)$$

where X is a vector of baseline characteristics and all other notation is as in (2). To test for balance, we conduct an F -test to test for joint significance of the coefficient vector β . We run separate versions of (2) in which X consists of student, teacher, or school-level characteristics.

Table 2 presents results. Of the 28 variables tested, 4 (S4 female enrollment, proportion of sampled students who are female, student employment, and student grit) differ significantly between treatment and control schools at the 5% level. In regressions of treatment status on groups of baseline variables, student characteristics are jointly significant at 5%. All baseline characteristics are jointly significant at 10%.¹⁸

Randomization fell under our control as researchers, making these differences simply bad luck. Moreover, the direction of any resulting bias is unclear: for instance, students in treated schools are grittier at baseline than those in control schools, but less likely to be employed. Bias from the greater prevalence of female students in treated schools can largely be mitigated by estimating results separately by gender. Nonetheless, we check robustness of our estimates of student outcomes by including a female dummy, baseline employment dummy, and baseline grit. Omitting these variables from the joint tests reported above restores baseline balance ($p=.16$ for student variables, $p=.25$ for all baseline characteristics).

III.4. Empirical Strategy

The main statistical method we use is ordinary least squares linear regression (OLS). This is the appropriate method as randomized control trials solve (in principle) the selection problem

¹⁸ There are some numerical discrepancies between results in Table 2 and those reported in the analogous table (Table 3) of Blimpo and Pugatch (2019). These discrepancies arise from two sources. First, we updated how missing values are coded in response to reviewer comments on Stage 1. Second, we omitted discussion and practice exercises from the definition of active instructional techniques, also in line with Stage 1. Baseline balance remains qualitatively unchanged.

for estimation of the mean outcome difference due to assignment to treatment. Moreover, OLS allows us to adjust easily for the stratification of treatment, ensuring that we rely on experimental variation. We cluster standard errors by school to account for correlated outcomes among students within a school. We report statistical significance in two ways. First, we use conventional p -values, with one, two, or three stars denoting significance at 1%, 5%, and 10%, respectively. Second, we report q -values following the Benjamini, Krieger, and Yekutieli (2006) procedure to control the false discovery rate (FDR) when conducting multiple hypotheses.¹⁹ This q -value is the analogue of the conventional p -value when controlling the FDR at level q . In other words, a q -value of 0.20 means the null hypothesis would be rejected only when permitting the FDR to be no less than 20%.

Our main estimating equation is the following regression:

$$y_{isg} = \alpha + \beta T_{sg} + \delta y_{0isg} + \gamma_g + \varepsilon_{isg} \quad (4)$$

where y is an outcome (with y_0 the outcome at baseline) and all other notation is as in (1). The coefficient of interest is β , which measures the intent to treat (ITT), or the effect of the offer of teacher training T on the mean outcome.²⁰ We use equation (4) to test hypotheses along each step of the program's theory of change. Outcomes specific to each hypothesis, including details on where these outcomes appear in the endline data, appear in the Stage 1 Registered Report (Blimpo and Pugatch 2019).

Where necessary, we supplement estimates of the intent to treat in (4) with instrumental variable estimation to account for partial compliance.

$$D_{isg} = \mu + \theta T_{sg} + \delta y_{0isg} + \gamma_g + v_{isg} \quad (5)$$

$$y_{isg} = \alpha + \rho D_{sg} + \delta y_{0isg} + \gamma_g + \varepsilon_{isg} \quad (6)$$

¹⁹ We calculate q -values within hypotheses defined in the Stage 1 Registered Report.

²⁰ To construct Lee bounds on the ITT from equation (4), we proceed in two steps. First, we regress the outcome on all right-hand side variables in (4) except treatment. Second, we regress the residuals from step one on treatment. We calculate Lee bounds on the coefficient in the second step. This procedure differs from the Lee (2009) method to calculate bounds in the presence of covariates because some covariate-by-strata cells lack variation in treatment. In some cases, the bounds we calculate fail to contain the point estimate due to spurious correlation between included covariates and treatment assignment, and therefore should be interpreted with caution.

in which D is an indicator for whether student i 's teacher attended training and T is an indicator for assignment to the treatment group. Equation (5) is the first stage equation for treatment take-up, while (6) is the second stage, in which T instruments for D . The parameter ρ measures the local average treatment effect (LATE), or the effect of the program for students whose teachers complied with their experimental assignment.

Our compliance measures are training attendance and exchange visit attendance. For each measure of compliance, the attendance variable (D) takes a value of one if at least one teacher from the school attended the training session, zero otherwise.²¹ We then measure compliance as the school-level mean of this variable. Following the analysis plan, we estimate the LATE for any compliance measure falling below 85%.

IV. Results

IV.1. Main findings

IV.1.1. Uptake and implementation: do schools adhere to the in-service training?

Schools and teachers adhere to the program, as shown by the take-up measures reported in Table 3. Columns 1-2 reports, respectively, the impact on the proportions i) of trainings attended by at least one teacher from the school; ii) of exchange visits attended by at least one teacher from the school.²² Control group attendance was negligible (2%), with contamination driven by the confusion in initial treatment assignment mentioned previously. The coefficients on treatment in these columns therefore approximate take-up rates. Overall, training attendance was 88% among treated schools, while exchange visit attendance was 81%. In other words, treated schools sent teachers to more than five of six training sessions, on average, and more than four of every five exchange visits.

²¹ We define take-up at the school level using administrative data. Defining take-up as attendance by the teacher surveyed at baseline would be problematic. We surveyed S4 teachers at baseline, but at 71 percent of schools a different teacher(s) will teach entrepreneurship to the study cohort in S5-S6. Ideally, we would measure take-up by matching the study cohort to their current entrepreneurship teacher and tracking the teacher's take-up in that year. Unfortunately, the administrative records fail to report the grade(s) taught by the teacher, or whether the teacher teaches the particular group of students sampled at baseline. In the absence of this information, we find it simplest to define take-up at the school level.

²² We do not have compliance data from the support visits. However, we expect non-compliance with support visits to be very low. Youth Leaders (the NGO staff who made support visits) were evaluated regularly on their performance. No Youth Leader was dismissed or disciplined for poor performance related to program delivery during the project period.

Because exchange visit attendance fell below 85%, we also report estimates of equation (6) for this measure of compliance throughout the paper. With a first-stage coefficient of 0.79 (Table 3, column 2), our LATE estimates scale the ITT by $1/0.79 = 1.27$.

High program uptake is also reflected in the classrooms through measures of curricular implementation (columns 3-5). Scheduling of Skill Labs, which is required by the reform for all schools, increased by 52 percentage points due to treatment, a six-fold increase over the control mean of 8%. We did not observe an impact on use of lesson plans, for which the base is somewhat high (61%). Teachers in the treatment group scored 8 percentage points, or 14%, higher on a test of the entrepreneurship curriculum. Trainings for the treatment group emphasize each of these outcomes.

We find no effect on administrators' perceptions about the reform (columns 6-7). Head teachers have somewhat mixed feelings about the student-centered teaching methods. On a scale from 0 (least supportive) to 6 (most supportive), principals reported an average score of 3.6 in the control group, not statistically discernable from the treatment group. This result contrasts with overwhelming support (93%) by head teachers in both groups that teachers' primary goal should be to develop student skills, the main aim of the reform. The apparent contradiction may reflect the short-term tensions that arise with the implementation of major reforms. Administrators support the reform's focus on student skills, but are less supportive of non-traditional means to promote these skills. Program exposure did not change these beliefs. Moreover, a post-endline qualitative investigation we commissioned reported perceived tension between practical components of the curriculum and examination criteria. This tension prompted teachers to focus less on practical skills for S6 students due to the urgency of national examinations (Anand 2019).

All results remain robust to multiple hypothesis corrections. The local average treatment effects reported in Table A5 show, as expected, larger effects for compliers than the ITT. Skills Lab scheduling increased eight-fold vs six-fold for the ITT and teachers' knowledge of the curriculum increased by 17% versus 14% for the ITT.

Taken together, we conclude that these are important signs of adherence to the in-service training provided by the intervention.

IV.1.2. Are teachers adjusting their pedagogy?

We observed moderate early changes in pedagogy through classroom observations and student reports of teacher practices. Table 4 reports the treatment effect on two measures drawn from endline classroom observations: i) proportion of class time in active instruction, based on the Stallings classroom observation instrument (J. Stallings 1977; J. A. Stallings and Mohlman 1988); and ii) use of instructional techniques specific to the entrepreneurship curriculum, such as role play and group discussion.²³ For each of these measures, we report outcomes for all observations and for Skills Labs only, as the latter promote active instruction most forcefully within the curriculum. Moreover, we report outcomes from the Stallings instrument for the full 52-minute classroom observation and split by first/second half, as the first half involves setting up the lab structure and other classroom management.

Treatment did not significantly alter active instruction time over the entirety of the classroom observation. This null effect conceals differences within the observation, however. Active instruction time fell 6 percentage points in the first half in response to treatment, but increased 10 percentage points in the second half. Both estimates have q -values of .05 or below. These results match the “build, practice, present” structure of the new curriculum, in which teachers are trained to explain an activity (a form of passive instruction) before providing students the opportunity to practice and present to the class.

Exploratory analysis looking more closely at the individual activities recorded in the Stallings observations is consistent with this interpretation (Table A15). Treated teachers spent 5 percentage points more time on lecture in the first half of the observation, but reduced lecturing by the same amount in the second half (column 1). Treatment also led to significantly reduced practice/drill and copying time by students in the second half, with a nearly commensurate increase in discussion time. In other words, treated classrooms became more active, particularly later in the class period.²⁴

Results change when restricting attention to Skills Labs (Table 4, columns 4-6). Treated teachers spend 18 percentage points *less* time in active instruction than control teachers in Skills Labs. The effect is driven by a 32 percentage point decrease in the second half, precisely when

²³ In the Stallings instrument, we define active instruction as the proportion of classroom time in Q&A/discussion, student presentation, and project/interactive activity. Active instructional techniques include group discussion, research, case study, role play, debate, finance/practice activity.

²⁴ We interpret these results with caution. In addition to being exploratory, the coefficients significant using conventional p -values are no longer significant when adjusted for multiple hypothesis tests.

instruction should be most active. What explains these puzzling results? Note that when restricting attention to Skills Labs, the sample size falls from 235 classroom observations to 70. Of these 70 observations, only 8 are from the control group. This small group of control teachers conducting Skills Labs is likely to be highly selected, with high motivation to follow the reformed curriculum.

Observed instructional techniques such as group discussion, research, case studies, debates, and role playing increased by 6 percentage points, or 19% relative to the control group (Column 7). Students reported increased use of active instructional techniques during regular class time by 3 and 7 percentage points, respectively, for a narrower measure capturing small group work, class presentations by students, small group presentations, and a broader measure adding outside class activities, group discussion, research, case studies, role playing, financial calculations, computer work, and career planning.

The statistical significance of all the results reported remains when we correct for multiple hypotheses testing. Table A6 reports the same results using instrumental variables to capture the local average treatment effects. The effect for compliers is 13 percentage points versus 10 for the ITT on active instruction time; 8 percentage points versus 6 for the ITT on the observed active instructional techniques; and 3 and 9 percentage points, respectively, for the narrower and broader student report of the use of active instructional techniques in regular classes.

Overall, the observed pedagogical changes are consistent with adherence to the reform and training.

IV.1.3. Has the intervention improved student learning outcomes?

We find no effect of the intervention on student academic and non-cognitive skills. Table 5 reports the impact on students' academic skills, as measured by performance on the S6 (grade 12) national entrepreneurship exam and the last available school-level promotional exam (columns 2-3).²⁵ We normalize scores to the control group mean and standard deviation. We find no effect on either of these measures. Important context here is that national exam content did not change dramatically under the 2016 curriculum reform. Exams continued to focus on mastery

²⁵ The Stage 1 Registered Report also included overall national exam score, but we were unable to obtain this data at the time of writing.

of curricular content as under the previous curriculum (Honeyman 2016), rather than demonstrate acquisition of practical skills. Individual schools may have altered promotional exams to align better with the new curriculum, but doing so would have been costly in the absence of a similar shift in national exams. Educate! reported no notable shifts in promotional exam content in treated schools. Such shifts in control schools would be even less likely.

Table 5, columns 4-9 capture students' financial and entrepreneurship skills, as proxied by patience, saving, or business knowledge.²⁶ We find some evidence that program exposure made students more patient, with a four percentage point increase in the proportion with an implied monthly discount rate below 300%. The result is significant at the 5% level, but does not survive the multiple hypothesis correction. There is no effect for an implied discount rate of 100% or less.²⁷ Conditional on students saving a positive amount, we find a negative effect of US\$4.40 on total savings, or about 20% of the control group mean. This effect may be due to savers in the treatment group investing in businesses. However, this effect is not significant at conventional levels after accounting for multiple hypothesis tests.

Table 6 reports no effects on students' non-cognitive skills, as measured by aspirations, locus of control, and grit. The results persist when we use instrumental variables estimation (Table A7).

IV.1.4. Do we observe more entrepreneurial activities and higher earnings?

The intervention increased students' participation in business activities. It also reduced the share employed, suggesting substitution between entrepreneurship and wage employment. Conditional on business participation, the treatment reduced business profits. Table 7 reports results. Column 2 reports a 5 percentage point increase in business participation, defined as earning money from running a business. This increase is 17% of the control group mean and is

²⁶ There are Skills Labs on savings, loans, and budgeting, which are intended to promote savings and forward-looking behavior. See Appendix 1.

²⁷ Discount rates are based on responses to the following endline survey questions:

- ESQ801: "Suppose someone you fully trust offered you 5,000RWF today. However, he tells you that you can wait one month and receive 10,000RWF instead. Which do you prefer?"
- ESQ802: "What if your choice was between 5,000RWF today or 20,000RWF one month from now?"

We code discount rate<100% if the respondent prefers 10,000RWF in question 801. We code discount rate<300% if the respondent prefers 20,000RWF in question 802.

significant at 5% (q -value=.09). The effect is driven by participation in student business clubs (3 percentage points, significant at 1%; column 4), a central component of the reformed curriculum. We find no effect on participation in family- or peer-run businesses (column 5). Taken together, the program is encouraging business formation, with the school as the main channel.

We find no effect of the program on the share of businesses outside the agricultural sector, nor on the share of businesses with employees (columns 6-7).

The program decreased the proportion of students in paid employment by 5 percentage points, significant at 1% (q =.03; column 8). The magnitude of the decrease is identical to the increase in business participation. Moreover, overall income was unaffected (column 9), suggesting substitution between these activities. These results are an early sign that the treatment successfully trained more entrepreneurs, a main program objective.

Conditional on business participation, we find a decrease of US\$15.80 on business profits (column 10). This effect is statistically significant and substantively large, representing 41% of the control mean. This negative effect on business profit may reflect lower business acumen by the new entrants. It may also reflect greater business investment by treated students, consistent with the lower savings reported in Table 5. However, exploratory analysis reveals no significant changes in business expenses in response to treatment (Table A16, columns 5-6).²⁸

In addition to business creation and income, the treatment also intended to promote more effective business practices among student entrepreneurs. Table A17 shows exploratory analysis for specific business practices as outcomes. Among all students, the treatment increased the proportion that had written business ideas and kept financial records. When conditioning on business participation, these increases are no longer statistically significant. However, treated businesses are less likely to report paying tax or registering with government, further evidence that businesses induced by the intervention are negatively selected. These results are merely suggestive, however, because the magnitudes are small (2-3 percentage points) and not statistically significant after adjusting for multiple hypotheses.

²⁸ Additional exploratory analysis from Table A16 shows that the decline in business profits in response to treatment is significant only when conditioning on business involvement (column 3), not when imputing zero profits to students not involved in business (column 4). In other words, the intervention did not reduce business profit overall. The table also shows that the decline in savings observed in Table 6 persists in the unconditional measure, i.e., imputing zero for those not reporting any savings (Table A16, column 2). Correcting for multiple tests makes these results less precise, however.

One possible outcome of this involvement in business activity while in school is that students drop out. For instance, a school-based financial literacy program in Ghana led to an increase in child labor, as students exposed to the program entered the labor market (Berry, Karlan, and Pradhan 2018). More generally, the economic opportunities available to youth influence their schooling decisions (e.g., Heath and Mobarak 2015; Atkin 2016; Pugatch 2018). In our case, we find no effect on dropout (Column 1).

For the same set of business outcomes, effects on compliers are larger but of similar magnitude, as expected (Table 12A). The effect on business participation is 6 percentage points, as is the effect on employment (compared to 5 percentage points for the ITT). Among those running a business, profit dropped 52% for compliers compared to 41% for the ITT estimates.

IV.2. Heterogeneity and Mechanisms

IV.2.1. Heterogeneous treatment effects

The effect of the program may differ among students or teachers. While the experiment was not designed to fully capture such heterogeneities, testing them can nevertheless provide important cues for policy nuances and program adjustments. We allow treatment effects to vary according to observable characteristics of a student or teacher by modifying (4) as:

$$y_{isg} = \alpha + \beta_1 T_{sg} + \beta_2 (T_{sg} \times X_{0isg}) + \beta_3 X_{0isg} + \delta y_{0isg} + \gamma_g + \varepsilon_{isg} \quad (7)$$

where X_0 is some student and teacher characteristic determined prior to the treatment and pre-specified in our analysis plan. A non-zero value of β_2 indicates that the effect of treatment differs according to X_0 .

We limit estimation of equation (7) to student business involvement, employment, and total income from business and employment. Table 8 reports results. In columns 1-6, X_0 is (respectively) student gender, past academic performance (S3 exam score), household socioeconomic status (the first principal component of parents' education, household assets, and parents' occupation), and teachers' gender, years of teaching experience, and qualification status.

We find three main results. First, student business involvement is negatively mediated by teacher qualification. Put another way, students taught in treatment schools increase business participation by 11 percentage points, but this increase is reversed entirely for students of

qualified teachers. This result may suggest resistance to change among qualified teachers, whereas unqualified teachers may be more receptive to new ideas and methods. Alternatively, qualified teachers may focus more on academic outcomes instead of encouraging students to form businesses while in school.

Second, the intervention has a differential positive effect on the employment status of high performing students. The results suggest that there may be a negative sorting, whereby poor performers are more likely to try business or entrepreneurial activities, whereas high performers are able to obtain jobs. The finding is consistent with previous literature suggesting that a large share of entrepreneurs in Africa become entrepreneurs or self-employed by necessity (e.g., Kingdon, Sandefur, and Teal 2006; Günther and Launov 2012; Fox, Senbet, and Simbanegavi 2016). Finally, students whose family's socio-economic status is above median earn less income (significant at 10%), which may reflect lower need to engage in economic activities while in school.

Which students are creating businesses in response to the intervention? Is business creation driven by those who would likely have created businesses anyway? Or did the intervention generate unlikely entrepreneurs? In exploratory analysis, we test whether treatment effects differ by prior likelihood of starting a business. First, we predict each student's propensity to start a business, using parameters estimated among the control group (via logit regression of endline business participation on all baseline characteristics in Table 2, plus baseline business ownership and randomization strata). We then interact treatment status with quartiles of this propensity score. We also include main effects of each quartile in the regression.

Table A18 presents results. We find that students most likely to start businesses (quartile 4) are 6 percentage points less likely to drop out of school in response to treatment, in comparison with control group students from the same quartile (column 1). This finding suggests that likely entrepreneurs found the program sufficiently valuable to remain in school. Overall, business participation in response to treatment was not driven by any particular quartile (column 2). However, increases in student business club activity were driven by students least likely to create businesses (i.e., the first quartile; column 4). This result suggests that student business clubs were an effective vehicle to draw otherwise unlikely entrepreneurs into business, though the effect size is modest (2 percentage points). Students from the top quartile were less likely to start businesses outside the agricultural sector, a puzzling result (column 6).

Finally, declines in wage employment in response to treatment were driven by students from the top 2 quartiles of likely business creation (column 8). Yet these same quartiles did not see commensurate increases in business creation (column 2), suggesting that the apparent substitution between entrepreneurship and employment in treated schools occurred across students with different propensities to participate in business. We interpret these results with caution, however, as significance levels are mostly 10% using conventional p-values, and no longer significant when adjusting for multiple hypothesis tests.

Because male and female students face different challenges in completing secondary school and transitioning to economic activity, we assess the impact on all student outcomes separately by student gender. Tables A2 through A4 report results. We find a negative treatment effect on the entrepreneurship exam of 0.1 standard deviations for boys, significant at 10%. The point estimate is of equal magnitude but positive for girls, though not statistically significant. On the other hand, the increase in patience previously reported (i.e., a monthly discount rate below 300%) is entirely driven by boys. We found no impact on saving for the entire sample, but the split sample shows a 5 percentage point increase in the share of girls who save (significant at 10 percent level). Moreover, decreases in the amount saved (conditional on any savings) is statistically significant only for girls. This result may reflect lower savings among new savers, thus reducing the average in the treatment group. We find no treatment effects on non-cognitive skills for boys or girls (Table A3).

In term of economic activities (Table A4), overall business participation is significantly positive only for boys (at 10%), whereas student business club participation is significantly positive for both boys and girls. Business clubs are thus a promising avenue to involve more girls in business activities. Another notable difference is that the negative effect on business profits (conditional on business participation) is largely driven by boys. The point estimate is negative, smaller, and not statistically significant for girls.

IV.2.2. Mechanisms

To learn about aspects of the intervention that were most or least influential, we ran a regression analogous to equation (7), where we replace X with M , a hypothesized mechanism through which the intervention influences outcomes:

$$y_{isg} = \alpha + \beta_1 T_{sg} + \beta_2 (T_{sg} \times M_{isg}) + \beta_3 M_{isg} + \delta y_{0isg} + \gamma_g + \varepsilon_{isg} \quad (8)$$

A non-zero value of β_2 in equation (8) now represents a differential effect of the program according to values of M . For instance, if M measures active instruction, then $\beta_2 > 0$ indicates that students whose teachers used active instruction more intensively increased their outcomes more than other students in the program.²⁹

We test two types of mechanisms, the take-up of program elements (trainings and exchange visits) and pedagogy (three measures of active instruction). Table 9 shows results. We find no differential effects of the program on business involvement, employment, or earnings according to any of the hypothesized mechanisms.

V. Conclusion

We evaluate an intensive, two-year in-service teacher training and support program among secondary school entrepreneurship teachers in Rwanda. We find mixed results. On the one hand, high uptake among teachers, demonstrable pedagogical change, and increased student engagement in business activities are positive early signs that the program may be on track to achieve longer-term goals. On the other hand, pedagogy did not change in all ways envisioned by the program. Nor did the program translate into improved student learning outcomes or robust increases in skills, raising several questions with respect to the theory of change.

One puzzle is the lack of impact on student learning. Our preferred explanation is a mismatch between the program and curricular focus on practical skills and the certification exams' focus on content knowledge. There are however, other potential explanations. For example, teachers' adaptation to the new approach may have substituted for exam preparation, leading to no short-term improvements in learning. It is also possible that, for learning outcomes, the default training is as effective as the training and support delivered to the treatment group. These results underscore the complexity involved to change pedagogy and improve learning.

²⁹ This approach is worthwhile but faces two major drawbacks. First, M is an intermediate outcome of the program, i.e., the program alters y through its effect on M . It is therefore not entirely clear how to interpret a program effect that holds M constant; Angrist and Pischke (2008) refer to this as the problem of "bad control." Second, M is not randomly assigned among teachers. If M is correlated with unobserved teacher attributes that also affect the outcome (such as motivation), then β_2 will be a biased estimate of M 's role as a treatment effect mechanism. These caveats should be kept in mind when interpreting results.

Full assessment of program impact must also account for its costs. The intervention cost US\$71 per student over the three years of the program, or an annual average of US\$24 per student.³⁰ The program increased business participation by 5 percentage points. In other words, an average of 20 student beneficiaries was induced to generate one additional business at endline relative to the control group. The cost of generating one additional business was therefore US\$71 x 20 = US\$1,420, or US\$473 per year. This increase in business participation was exactly offset by a reduction in wage employment. Nor was there any impact on businesses with paid employees. Therefore, these costs were incurred without netting any new jobs.

Nonetheless, these results are short-term, with outcomes measured before students had completed secondary school. A follow-up study after the cohort of interest enters the labor market is necessary to understand fully the policy implications of the program. We plan future tracer studies to inform these questions.

³⁰ Calculation based on total program cost of US\$4,730 and mean of 67 students in entering cohort per treated school at baseline. Cost include program administration, targeting, staff and user training, implementation, user costs, averted costs, and monitoring, following the J-PAL cost effectiveness analysis template.

References

- Abeberese, Ama Baafra, Todd J. Kumler, and Leigh L. Linden. 2014. "Improving Reading Skills by Encouraging Children to Read in School: A Randomized Evaluation of the Sa Aklat Sisikat Reading Program in the Philippines." *Journal of Human Resources* 49 (3): 611–633.
- African Economic Outlook. 2016. "Promoting Youth Employment." 2016. http://www.africaneconomicoutlook.org/en/theme/youth_employment/.
- Alaref, Jumana Jamal Subhi, Stefanie Brodmann, and Patrick Premand. 2019. "The Medium-Term Impact of Entrepreneurship Education on Labor Market Outcomes: Experimental Evidence from University Graduates in Tunisia." SSRN Scholarly Paper ID 3316478. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=3316478>.
- Albornoz, Facundo, Maria Victoria Anauati, Melina Gabriela Furman, Mariana Luzuriaga, Maria Eugenia Podesta, and Ines Taylor. 2018. "Training to Teach Science: Experimental Evidence from Argentina." SSRN Scholarly Paper ID 3252234. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=3252234>.
- Alibhai, Salman, Niklas Buehren, and Sreelakshmi Papineni. 2016. "From Learning to Earning."
- Anand, Gautam. 2019. "Entrepreneurship Education and Teacher Training in Rwanda: Findings of the Qualitative Research Study."
- Angrist, Joshua D., and Victor Lavy. 2001. "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools." *Journal of Labor Economics* 19 (2): 343–69. <http://www.jstor.org/action/showPublication?journalCode=jlaboreconomics>.
- Angrist, Joshua D., and Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. 1st ed. Princeton University Press.
- Atkin, David. 2016. "Endogenous Skill Acquisition and Export Manufacturing in Mexico." *The American Economic Review* 106 (8): 2046–2085.
- Awusu, Solomon, and Moussa P. Blimpo. 2019. "Human Capital." In *The Future of Work in Africa: Harnessing the Potential of Digital Technologies for All*, edited by Jieun Choi, Mark Dutz, and Zainab Usman, 72–106. Washington DC: World Bank.
- Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *The Quarterly Journal of Economics* 122 (3): 1235–64.
- Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. 2006. "Adaptive Linear Step-up Procedures That Control the False Discovery Rate." *Biometrika* 93 (3): 491–507. <https://doi.org/10.1093/biomet/93.3.491>.
- Berlinski, Samuel, and Matias Busso. 2017. "Challenges in Educational Reform: An Experiment on Active Learning in Mathematics." *Economics Letters* 156 (July): 172–75. <https://doi.org/10.1016/j.econlet.2017.05.007>.
- Berry, James, Dean Karlan, and Menno Pradhan. 2018. "The Impact of Financial Education for Youth in Ghana." *World Development* 102: 71–89.
- Beuermann, Diether W., Emma Naslund-Hadley, Inder J. Ruprah, and Jennelle Thompson. 2013. "The Pedagogy of Science and Environment: Experimental Evidence from Peru." *The Journal of Development Studies* 49 (5): 719–736.
- Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2014. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." *Quarterly Journal of Economics* 129 (2): 697–752. <http://qje.oxfordjournals.org/content/by/year>.
- . 2018. "The Long Term Impacts of Grants on Poverty: 9-Year Evidence from Uganda's Youth Opportunities Program." SSRN Scholarly Paper ID 3223028. Rochester, NY: Social Science Research Network. <https://papers.ssrn.com/abstract=3223028>.
- Blimpo, Moussa P., and Todd Pugatch. 2019. "Entrepreneurship Education and Teacher Training in Rwanda." *Journal of Development Economics* 140 (September): 186–202. <https://doi.org/10.1016/j.jdeveco.2019.05.006>.

- Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur. 2018. "Experimental Evidence on Scaling up Education Reforms in Kenya." *Journal of Public Economics* 168 (December): 1–20. <https://doi.org/10.1016/j.jpubeco.2018.08.007>.
- Bruns, Barbara, Leandro Costa, and Nina Cunha. 2018. "Through the Looking Glass: Can Classroom Observation and Coaching Improve Teacher Performance in Brazil?" *Economics of Education Review*, March. <https://doi.org/10.1016/j.econedurev.2018.03.003>.
- Bux, Sara. 2016. "The Effect of Entrepreneurship Education Programmes on the Mind-Set of South African Youth." PhD Thesis, University of Pretoria.
- Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas-Shady, and Yuri Soares. 2011. "The Labor Market Impacts of Youth Training in the Dominican Republic." *Journal of Labor Economics* 29 (2): 267–300.
- Cilliers, Jacobus, Brahm Fleisch, Cas Prinsloo, and Stephen Taylor. 2019. "How to Improve Teaching Practice? An Experimental Comparison of Centralized Training and in-Classroom Coaching." *Journal of Human Resources*, February, 0618-9538R1. <https://doi.org/10.3368/jhr.55.3.0618-9538R1>.
- Educate! 2014. "Impact Evaluation of Midline Data from the Educate! Randomized Control Trial."
- Evans, David K., and Anna Popova. 2016. "What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews." *World Bank Research Observer* 31 (2): 242–270.
- Filmer, Deon, and Louise Fox. 2014. *Youth Employment in Sub-Saharan Africa*. Africa Development Forum. The World Bank. <https://doi.org/10.1596/978-1-4648-0107-5>.
- Fox, Louise, Lemma W. Senbet, and Witness Simbanegavi. 2016. "Youth Employment in Sub-Saharan Africa: Challenges, Constraints and Opportunities." *Journal of African Economies* 25 (suppl_1): i3–15. <https://doi.org/10.1093/jae/ejv027>.
- Glewwe, P., and K. Muralidharan. 2016. "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications." *Handbook of the Economics of Education* 5.
- Glewwe, Paul, Eric A. Hanushek, Sarah Humpage, and Renato Ravina. 2013. "School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010." *Education Policy in Developing Countries* 4 (14,972): 13.
- Günther, Isabel, and Andrey Launov. 2012. "Informal Employment in Developing Countries: Opportunity or Last Resort?" *Journal of Development Economics* 97 (1): 88–98. <https://doi.org/10.1016/j.jdeveco.2011.01.001>.
- Haushofer, Johannes, and Jeremy Shapiro. 2015. "Household Response to Income Changes: Pre-Analysis Plan." <https://www.socialscisearch.org/docs/analysisplan/384/document>.
- Heath, Rachel, and A. Mushfiq Mobarak. 2015. "Manufacturing Growth and the Lives of Bangladeshi Women." *Journal of Development Economics* 115: 1–15.
- Hicks, Joan Hamory, Michael Kremer, Isaac Mbiti, and Edward Miguel. 2015. "Vocational Education in Kenya-A Randomized Evaluation." *3ie Grantee Final Report. New Delhi: International Initiative for Impact Evaluation (3ie)*.
- Honeyman, Catherine A. 2016. *The Orderly Entrepreneur: Youth, Education, and Governance in Rwanda*. 1 edition. Stanford, California: Stanford University Press.
- Huber, Laura Rosendahl, Randolph Sloof, and Mirjam Van Praag. 2014. "The Effect of Early Entrepreneurship Education: Evidence from a Field Experiment." *European Economic Review* 72 (November): 76–97. <http://www.sciencedirect.com/science/journal/00142921>.
- Kerwin, Jason T., and Rebecca Thornton. 2015. "Making the Grade: Understanding What Works for Teaching Literacy in Rural Uganda." *Unpublished Manuscript. University of Illinois, Urbana, IL*.
- Kingdon, Geeta, Justin Sandefur, and Francis Teal. 2006. "Labour Market Flexibility, Wages and Incomes in Sub-Saharan Africa in the 1990s." *African Development Review/Revue Africaine de Développement* 18 (3): 392–427. <https://doi.org/10.1111/%28ISSN%291467-8268/issues>.

- Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies* 76 (3): 1071–1102.
- Loyalka, Prashant, Anna Popova, Guirong Li, and Zhaolei Shi. 2019. "Does Teacher Training Actually Work? Evidence from a Large-Scale Randomized Evaluation of a National Teacher Training Program." *American Economic Journal: Applied Economics* 11 (3): 128–54. <https://doi.org/10.1257/app.20170226>.
- Lucas, Adrienne M., Patrick J. McEwan, Moses Ngware, and Moses Oketch. 2014. "Improving Early-Grade Literacy In East Africa: Experimental Evidence From Kenya And Uganda." *Journal of Policy Analysis and Management* 33 (4): 950–976.
- Mbiti, Isaac, Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, and Rakesh Rajani. 2019. "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania." *The Quarterly Journal of Economics* 134 (3): 1627–73. <https://doi.org/10.1093/qje/qjz010>.
- Murnane, Richard J., and Alejandro J. Ganimian. 2014. "Improving Educational Outcomes in Developing Countries: Lessons from Rigorous Evaluations." Working Paper 20284. National Bureau of Economic Research. <http://www.nber.org/papers/w20284>.
- Null, Claire, Clemencia Cosentino, Swetha Sridharan, and Laura Meyer. 2017. "Policies and Programs to Improve Secondary Education in Developing Countries: A Review of the Evidence." Mathematica Policy Research. <https://www.mathematica-mpr.com/our-publications-and-findings/publications/policies-and-programs-to-improve-secondary-education>.
- Piper, Benjamin, and Stephanie Simmons Zuilkowski. 2015. "Teacher Coaching in Kenya: Examining Instructional Support in Public and Nonformal Schools." *Teaching and Teacher Education* 47: 173–183.
- Popova, Anna, David K. Evans, Mary E. Breeding, and Violeta Arancibia. 2018. *Teacher Professional Development around the World: The Gap between Evidence and Practice*. The World Bank.
- Premand, Patrick, Stefanie Brodmann, Rita Almeida, Rebekka Grun, and Mahdi Barouni. 2012. "Entrepreneurship Training and Self-Employment among University Graduates: Evidence from a Randomized Trial in Tunisia." http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2012/12/03/000158349_20121203172559/Rendred/PDF/wps6285.pdf.
- Pugatch, Todd. 2018. "Bumpy Rides: School-to-Work Transitions in South Africa." *Labour* 32 (2): 205–242.
- Pugatch, Todd, and Moussa Blimpo. 2016. "Teacher Training and Entrepreneurship Education: Evidence from a Curriculum Reform in Rwanda." American Economic Association. <https://doi.org/10.1257/rct.1030-3.0>.
- Robb, Alicia, Alexandria Valerio, and Brent Parton. 2014. *Entrepreneurship Education and Training: Insights from Ghana, Kenya, and Mozambique*. The World Bank.
- Rwanda Ministry of Education. 2016. "2015 Education Statistical Yearbook." http://www.mineduc.gov.rw/fileadmin/user_upload/Amatangazo/2015%20Education%20Statistical%20YearbookF.pdf.
- Stallings, J. A., and G. G. Mohlman. 1988. "Classroom Observation Techniques." *Educational Research, Methodology, and Measurement: An International Handbook*, 469–474.
- Stallings, Jane. 1977. *Learning to Look: A Handbook on Classroom Observation and Teaching Models*. Wadsworth Publishing Company.
- UNESCO Institute for Statistics. 2018. "UIS Statistics." 2018. <http://data.uis.unesco.org/>.
- World Bank. 2015. "World Development Indicators." World Bank. <http://data.worldbank.org/data-catalog/world-development-indicators>.
- . 2018a. *World Development Report 2018: Learning to Realize Education's Promise*. World Bank Washington, DC.
- . 2018b. *World Development Report 2019: The Changing Nature of Work*. World Bank.

Appendices

Appendix 1: Entrepreneurship Syllabus

The key competencies expected at the end of each grade level are:

At the end of senior four (S4), the learner should be able to:

- Exhibit the behavioral qualities of an entrepreneur
- Make rational career choices in daily life
- Make plans to reach their personal goals
- Evaluate the need for laws in business operation
- Analyze the role of standards in business
- Examine key components of a market and the role of market research
- Analyze the importance of management in a business organization
- Evaluate short- and long-term capital for future investment
- Evaluate the services/products offered by financial institutions.

At the end of senior five (S5), the learner should be able to:

- Generate business ideas and take advantage of opportunities
- Make valid contracts and resolve conflicts in business operations
- Justify the need for taxes in the economy
- Evaluate the factors that lead to business growth
- Analyze the role of technology in businesses and daily life.
- Maintain good relations with people at the workplace through effective communication
- Demonstrate ability and knowledge of carrying out general office operations
- Record accounting transactions and manage finances responsibly
- Exercise rights and responsibilities as an employee and employer
- Lead a team in accomplishing a goal

At the end of senior six (S6), the learner should be able to:

- Prepare a business plan for an enterprise
- Develop an ethical understanding of the Rwandan customs system
- Establish an effective quality compliance system in business activities
- Evaluate the contribution of entrepreneurship towards socio-economic development
- Analyze the Environmental Impact Assessment (EIA) as a tool for prevention and control of environmental impacts caused by socio-economic development

Skills Labs S4

- Intro to Entrepreneurship Process
- Creativity, Innovation, Invention
- Entrepreneurship as a Career
- Skills and Qualities
- Setting Goals
- Business Legal Formation

- The Ps of Marketing
- Competitor Survey
- Marketing Materials
- Quality Management
- Business Organizational Chart
- Personnel Management
- Fundraising for Sources of Capital
- Exploring Savings and Loans
- Record Keeping

Skills Labs S5:

- Generating Business ideas & Opportunities
- Business Contracts
- Business Taxes
- Market Survey
- Business Growth Strategies
- Effective Communication
- Business Skills and Customer Relations
- Business Documents
- Job Description
- Budgeting
- Financial Fitness Plan
- Journals
- Double Entry Accounting
- Rights and Responsibilities of workers and employers
- Safety Precautions
- Leadership Styles
- Developing a Team
- Problem Solving
- Conformity Assessment in Business

Skills Labs S6

- Role of Entrepreneurship in Social Economic Development
- Negative Effects of Economic Activity on the Environment
- EIA report
- Customs Procedures
- Importation and Exportation of Goods and Services in Rwanda
- Profit and Loss Account
- Balance Sheet
- Stock Control
- Marketing Plan
- Production Plan
- Business Plan
- Business Pitch
- Application of Metrology in Business Activities
- Writing a CV and application letter

- Interview Techniques

Tables

Table 1: Sample sizes and attrition

<u>Panel A: Sample sizes</u>				
	<u>Baseline</u>		<u>Endline</u>	
	<u>control</u>	<u>treatment</u>	<u>control</u>	<u>treatment</u>
	(1)	(2)	(3)	(4)
Schools	104	103	104	102
Teachers	104	103	166	167
<i>of which:</i>				
baseline	104	103	74	69
added sample			92	98
Students	1,554	1,541	1,447	1,433
<u>Panel B: Attrition from endline</u>				
	<u>control</u>	<u>treatment</u>	<u>difference</u>	<u>p-value</u>
	(1)	(2)	(3)	(4)
teachers	0.29	0.33	-0.04	0.60
students	0.07	0.07	0.00	0.95

Panel A shows baseline and endline sample sizes by treatment status. Schools considered in sample if at least one student who attended school at baseline completed survey. Panel B: teacher attrition=1 if teacher surveyed at baseline not surveyed at endline. Note that in cases where the current entrepreneurship teacher of the study cohort differed from the baseline teacher, there was sometimes insufficient time to survey the baseline teacher. p-values adjust for randomization strata.

Table 2: Baseline balance

Variable	control (1)	treatment (2)	difference (1) vs. (2)	p-value (1) vs. (2)
<u>School characteristics</u>				
boarding school	0.25 (0.04)	0.30 (0.05)	-0.05 (0.06)	0.25
S4 enrollment, male	24.9 (2.6)	22.2 (3.1)	2.7 (4.0)	0.39
S4 enrollment, female	32.9 (2.3)	41.9 (3.3)	-9.0 (4.0)	0.02
teachers, upper secondary	12.5 (1.0)	13.1 (0.6)	-0.7 (1.1)	0.39
teacher absence (%), last 3 Tuesdays	0.05 (0.01)	0.05 (0.01)	0.00 (0.01)	0.96
currently has electricity	0.85 (0.04)	0.80 (0.04)	0.05 (0.05)	0.57
head teacher knows Skills Lab definition	0.06 (0.02)	0.07 (0.03)	-0.01 (0.03)	0.74
head teacher considers interactive pedagogies as effective	0.59 (0.05)	0.61 (0.05)	-0.03 (0.07)	0.75
<u>Teacher characteristics</u>				
female	0.36 (0.05)	0.38 (0.05)	-0.02 (0.07)	0.88
age	33.6 (0.7)	33.7 (0.7)	-0.1 (0.9)	0.85
qualified	0.60 (0.05)	0.66 (0.05)	-0.06 (0.07)	0.42
showed written entrepreneurship lesson plan	0.46 (0.05)	0.44 (0.05)	0.03 (0.07)	0.83
comfortable with interactive pedagogies	0.64 (0.05)	0.68 (0.05)	-0.04 (0.07)	0.41
can calculate business profit	0.92 (0.03)	0.92 (0.03)	0.00 (0.04)	0.84
knows definition of business profit	0.93 (0.03)	0.85 (0.04)	0.08 (0.04)	0.08
holds another job	0.25 (0.04)	0.35 (0.05)	-0.10 (0.06)	0.06
<u>Student characteristics</u>				
female	0.54 (0.02)	0.62 (0.02)	-0.09 (0.03)	0.00
household assets	0.29 (0.01)	0.29 (0.01)	0.00 (0.01)	0.82
mother completed primary school	0.54 (0.02)	0.57 (0.02)	-0.03 (0.03)	0.10
repeating S4	0.04 (0.01)	0.04 (0.01)	-0.01 (0.01)	0.33
S3 exam score (aggregate)	53.5 (0.8)	52.7 (1.0)	0.7 (1.3)	0.56
employed during school holiday	0.29 (0.02)	0.25 (0.02)	0.05 (0.02)	0.03
understands compound interest	0.68 (0.01)	0.64 (0.02)	0.04 (0.02)	0.11
has savings	0.33 (0.02)	0.30 (0.02)	0.03 (0.03)	0.22
can calculate business profit	0.55 (0.02)	0.51 (0.02)	0.04 (0.03)	0.21
wants to enroll in post-secondary	0.72 (0.02)	0.74 (0.02)	-0.02 (0.03)	0.36

plans to start a business	0.76 (0.01)	0.78 (0.01)	-0.01 (0.02)	0.28
grit index	2.88 (0.03)	3.01 (0.03)	-0.12 (0.05)	0.01
Schools	104	103		
Teachers	104	103		
Students	1,554	1,541		
<u>p-values from omnibus tests:</u>				
all school characteristics				0.62
all teacher characteristics				0.28
all student characteristics				0.00
all variables				0.07

Sample is baseline survey, conducted February-March 2016. Columns (1)-(2) show means by treatment status. Column (3) shows difference between (1) and (2). Column (4) shows p-value of difference, adjusted for stratification by district and public/non-public school. Standard errors in parentheses. Head teacher coded as considering interactive pedagogies to be effective if he/she lists two interactive methods (question and answer; group work; games; activities outside classroom; experiment; portfolio) as among three most effective. Teacher coded as comfortable with interactive pedagogies if he/she lists two interactive methods as among three most with which he/she is most comfortable. Household asset index is proportion of items owned among radio, television, telephone, refrigerator, bicycle, motorcycle, and automobile. Grit index is mean response on 1-5 scale (1=least, 5=most) to four items about passion and perseverance in pursuit of goals. Omnibus tests regress treatment on all baseline characteristics within indicated group. Omnibus test for all characteristics conducted at school level, with school-level means used for student characteristics. Missing values imputed to control group mean and dummies for missing included in all omnibus tests.

Table 3: Program take-up and curricular implementation

	<u>take-up</u>		<u>teacher curricular implementation</u>			<u>administrator perceptions</u>	
	<u>training attendance</u>	<u>exchange visit attendance</u>	<u>Skills Lab scheduled</u>	<u>lesson plans</u>	<u>entrepreneurship knowledge</u>	<u>student-centered teaching</u>	<u>skill-based learning</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
treatment	0.85 (0.03)***	0.79 (0.03)***	0.52 (0.06)***	-0.04 (0.05)	0.08 (0.02)***	0.16 (0.15)	0.03 (0.03)
N	207	207	235	333	333	204	204
R-squared	0.86	0.85	0.39	0.07	0.14	0.16	0.07
Control mean	0.02	0.02	0.08	0.61	0.58	3.63	0.93
Baseline mean	N/A	N/A	N/A	0.59	N/A	3.25	N/A
q-value	0.00	0.00	0.00	0.22	0.00	0.22	0.22
Lee bounds	N/A	N/A	[0.47,0.50]	[-0.05,-0.03]	[0.06,0.08]	[0.10,0.19]	[0.02,0.03]

Table shows estimates of equation (4), including baseline outcome where indicated. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Training and exchange visit attendance (columns 1-2) measure proportion of teacher training or exchange visits attended by any teacher from school (of 6 total), 2016-2018. Skills Lab scheduled (column 3) is an indicator for whether the teacher reports Skills Lab in its weekly schedule when enumerators called to plan their endline school visit. Baseline outcome for Skills Lab scheduled is knowledge of Skills Lab definition. Lesson plans (column 4) is mean of indicators for showing entrepreneurship lesson plan and notes. Entrepreneurship knowledge (column 5) is mean of questions about entrepreneurship curriculum (baseline outcome included in analysis, but baseline mean not reported because question set differs). Column 6 is index of head teacher support for student-centered teaching methods (0=least supporting, 6=most supportive). Student-centered pedagogical tools include question and answer; group work; games; activities outside classroom; experiment; portfolio. Column 7 is indicator for whether head teacher agrees that teachers' primary goal should be to develop student skills. Baseline outcome for column 7 is mean of questions on support for skill-based learning (BHQ606/608/610).

Table 4: Pedagogical change

	active instruction time						active instructional techniques			
	all observed classes			Skills Labs only			observed		student reports	
	full observation	first half	second half	full observation	first half	second half	all classes	Skills Labs only	narrow	broad
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
treatment	0.01 (0.03)	-0.06 (0.03)*	0.10 (0.04)**	-0.18 (0.08)**	-0.07 (0.09)	-0.32 (0.14)**	0.06 (0.02)**	-0.03 (0.11)	0.03 (0.01)***	0.07 (0.01)***
N	235	235	235	70	70	70	235	70	2,611	2,611
R-squared	0.12	0.11	0.12	0.41	0.40	0.24	0.13	0.32	0.02	0.06
Control mean	0.48	0.45	0.52	0.64	0.52	0.78	0.31	0.31	0.95	0.62
q-value	0.30	0.05	0.04	0.04	0.19	0.04	0.03	0.31	0.00	0.00
Lee bounds	[0.01,0.02]	[-0.06,0.05]	[0.09,0.10]	[-0.13,-0.08]	[-0.07,0.01]	[-0.23,-0.13]	[0.05,0.06]	[-0.05,0.01]	[0.02,0.03]	[0.07,0.07]

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parenthesis (robust for columns 1-8, clustered by school for columns 9-10). * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Active instruction time is the proportion of classroom time in Q&A/discussion, student presentation, and project/interactive activity. The baseline outcome for columns (1)-(6) is an index of support for interactive teaching. Defined as $\text{index} = \text{BTQ300} + (3 - \text{BTQ301}) + \text{BTQ302}$. BTQ301 enters negatively because it is number of interactive techniques among least comfortable for teacher. Scale of raw index is 0-12 (least to most interactive). Active instructional techniques observed (columns 7-8) include group discussion, research, case study, role play, debate, finance/practice activity. Active instructional techniques from student reports: "narrow" measure in column 8 is mean of indicators for at least one use of small group work, presentation to class, and small group presentation to class. "Broad" measure in column 10 is mean of all indicators from narrow measure, plus outside class activity, group discussion, research, case study, role play, financial calculations, computer work, and career planning.

Table 5: Student academic outcomes and entrepreneurship skills

	Academic skills			Entrepreneurship skills					
	exam scores			monthly		understands	savings		entrepreneurship
	overall	S6	S4/S5	discount rate	discount rate	compound	any	amount	knowledge
(1)	entrepreneurship	promotional	<100%	<300%	interest	(7)	(if >0)	(9)	
treatment	N/A	-0.01 (0.05)	-0.3 (0.5)	0.01 (0.02)	0.05 (0.02)**	0.00 (0.02)	0.02 (0.02)	-5.4 (1.6)***	0.00 (0.01)
N		2,361	2,607	2,879	2,877	2,880	2,880	1,828	2,880
R-squared		0.44	0.27	0.03	0.04	0.02	0.05	0.10	0.04
Control mean		0.00	61.6	0.14	0.32	0.81	0.63	27.6	0.69
Baseline mean		-0.13	52.5	0.24	0.40	0.65	0.32	N/A	0.66
q-value		1.00	1.00	1.00	0.11	1.00	1.00	0.02	1.00
Lee bounds		[-0.017,-0.002]	[-0.33,-0.19]	[-0.003,0.001]	[-0.001,0.001]	[-0.001,0.000]	[-0.001,0.001]	[-0.04,0.11]	[-0.001,0.001]

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Exam scores normalized to control mean and standard deviation. Baseline outcome for academic outcomes (columns 1-3) is S3 exam score. Baseline outcome for column (8) is dummies for savings between RWF1-4,999, RWF5,000-10,000, or more than RWF10,000 (not reported in table because this is a vector). Monthly discount rate based on stated preference for 5,000RWF today versus larger amount one month from now. Savings in column (8) reported in USD using exchange rate on July 1, 2018, winsorized at 99th percentile. Entrepreneurship knowledge is proportion correct on 7-item test.

Table 6: Student non-cognitive skills

	<u>university or beyond</u> (1)	<u>Aspirations business or professional</u> (2)	<u>business creation</u> (3)	<u>Locus of control</u> (4)	<u>Grit</u> (5)
treatment	0.01 (0.02)	0.02 (0.02)	-0.01 (0.01)	-0.03 (0.06)	-0.03 (0.04)
N	2,880	2,880	2,880	2,880	2,880
R-squared	0.05	0.03	0.02	0.04	0.05
Control mean	0.80	0.66	0.98	6.39	3.37
Baseline mean	0.73	0.60	0.77	5.33	2.94
q-value	1.00	1.00	0.71	1.00	1.00
Lee bounds	[0.004,0.007]	[0.013,0.015]	[-0.01,-0.009]	[-0.04,-0.027]	[-0.029,-0.024]

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Locus of control is mean of 5 items about personal control over outcomes, with each item scaled from 1 (no control) to 10 (total control). Grit is mean of 4 items about personal persistence, scaled from 1 (not true) to 5 (very true).

Table 7: Student economic activity

	<u>dropped</u>		<u>business participation</u>			<u>business characteristics</u>		<u>employment</u>	<u>income</u>	
	<u>out</u>	<u>all</u>	<u>own</u>	<u>student</u>	<u>family/</u>	<u>non-</u>	<u>has paid</u>		<u>total</u>	<u>business</u>
	(1)	(2)	(3)	<u>club</u>	<u>peers</u>	<u>agricultural</u>	<u>employees</u>	(8)	(9)	<u>profit</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
treatment	0.00	0.05	0.03	0.03	0.00	-0.01	0.01	-0.05	-0.8	-15.8
	(0.01)	(0.02)**	(0.02)	(0.01)***	(0.01)	(0.01)	(0.01)	(0.02)***	(0.7)	(6.1)**
N	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,869	789
R-squared	0.02	0.06	0.06	0.02	0.01	0.03	0.02	0.02	0.07	0.05
Control mean	0.09	0.30	0.27	0.01	0.04	0.09	0.07	0.19	8.7	38.2
Baseline mean	0.00	0.22	0.22	0.00	0.09	0.04	0.22	0.13	1.5	4.7
q-value	0.68	0.09	0.28	0.00	0.63	0.63	0.63	0.03	0.35	0.03
Lee bounds	[0.003,0.004]	[0.043,0.045]	[0.026,0.028]	[0.029,0.03]	[0.003,0.005]	[-0.007,-0.005]	[0.005,0.006]	[-0.049,-0.047]	[-0.80,-0.59]	[-20.9,-13.5]

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Business involvement (all) is indicator for choice (a) or (c) in item ESQ401. Business characteristics refer to main business only. Baseline outcome for "business has employees" is indicator for involvement in any business. Employment is indicator for choice (b) or © in item BSQ401. Income measured in USD (using exchange rate from July 1, 2018), previous two months. Total income (column 9) uses self-reported total. Business profit (column 10) conditions on business involvement. Business profit adjusted by reported frequency of business earnings to estimate profits over two months. Alternative measure of total income is sum of reported employment income and business, adjusted by reported frequencies to estimate income for two months. All financial variables winsorized at 99th percentile.

Table 8: Heterogeneous treatment effects

	<u>Interaction term</u>					
	<u>female student</u>	<u>baseline exam score</u>	<u>above median SES</u>	<u>female teacher</u>	<u>teacher experience</u>	<u>qualified teacher</u>
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Outcome = business involvement [ESQ401/BSQ600]						
treatment	0.06 (0.03)*	0.05 (0.02)**	0.06 (0.03)**	0.05 (0.03)*	0.02 (0.04)	0.11 (0.04)***
treatment x interaction	-0.02 (0.04)	0.02 (0.02)	-0.03 (0.04)	-0.06 (0.05)	0.00 (0.00)	-0.13 (0.05)**
N	2,880	2,779	2,880	2,487	2,487	2,487
R-squared	0.07	0.07	0.06	0.06	0.06	0.07
Control mean	0.30	0.30	0.30	0.30	0.30	0.30
q value (treatment)	0.22	0.11	0.13	0.28	0.73	0.06
q value (interaction)	0.73	0.41	0.59	0.41	0.77	0.11
Panel B: Outcome = employment [ESQ401/BSQ500]						
treatment	-0.04 (0.03)	-0.05 (0.02)***	-0.03 (0.02)	-0.07 (0.02)***	-0.07 (0.03)***	-0.04 (0.03)
treatment x interaction	0.01 (0.03)	0.04 (0.01)***	-0.04 (0.03)	0.02 (0.04)	0.00 (0.00)	-0.04 (0.04)
N	2,880	2,779	2,880	2,487	2,487	2,487
R-squared	0.04	0.03	0.03	0.03	0.03	0.03
Control mean	0.19	0.19	0.19	0.19	0.19	0.19
q value (treatment)	0.37	0.06	0.37	0.06	0.06	0.41
q value (interaction)	0.80	0.06	0.41	0.73	0.59	0.60
Panel C: Outcome = income [ESQ401a/BSQ503]						
treatment	0.13 (1.37)	-0.74 (0.70)	0.28 (0.89)	-1.55 (1.03)	-1.73 (1.22)	-0.23 (1.19)
treatment x interaction	-0.59 (1.42)	0.04 (0.64)	-2.14 (1.16)*	1.48 (1.57)	0.11 (0.13)	-1.21 (1.54)
N	2,869	2,768	2,869	2,480	2,480	2,480
R-squared	0.10	0.07	0.07	0.07	0.07	0.07
Control mean	8.7	8.6	8.7	8.5	8.5	8.5
q value (treatment)	0.81	0.59	0.73	0.37	0.37	0.77
q value (interaction)	0.73	0.81	0.22	0.59	0.60	0.60

Table shows estimates of equation (4), including baseline outcome and main effect of term interacted with treatment. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Income measured in USD (using exchange rate from July 1, 2018), previous two months. All financial variables winsorized at 99th percentile. Baseline exam score normalized to mean zero and standard deviation one. SES is first principal component of household assets, parents' education, and indicator for parents in business or professional occupation. All interaction terms measured at baseline.

Table 9: Mechanisms

	Interaction term				
	<u>proportion</u> <u>trainings</u> <u>attended</u>	<u>proportion</u> <u>exchange</u> <u>visits</u> <u>attended</u>	<u>active</u> <u>instruction</u> <u>time</u>	<u>active</u> <u>instructional</u> <u>techniques</u> <u>(observed)</u>	<u>active</u> <u>instructional</u> <u>techniques</u> <u>(student reports)</u>
	(1)	(2)	(3)	(4)	(5)
Panel A: Outcome = business involvement [ESQ401/BSQ600]					
treatment	0.10 (0.07)	0.10 (0.07)	0.13 (0.07)*	0.13 (0.07)*	0.16 (0.13)
treatment x interaction	-0.03 (0.09)	-0.01 (0.11)	-0.12 (0.15)	-0.15 (0.18)	-0.12 (0.14)
N	2,880	2,880	1,926	1,926	2,611
R-squared	0.06	0.06	0.08	0.08	0.07
Control mean	0.30	0.30	0.30	0.30	0.30
q value (treatment)	0.83	0.83	0.70	0.70	0.96
q value (interaction)	1.00	1.00	1.00	1.00	1.00
Panel B: Outcome = employment [ESQ401/BSQ500]					
treatment	-0.12 (0.03)***	-0.09 (0.04)**	-0.12 (0.06)**	-0.08 (0.06)	-0.17 (0.10)
treatment x interaction	0.16 (0.11)	0.15 (0.11)	0.11 (0.11)	0.03 (0.16)	0.12 (0.11)
N	2,880	2,880	1,926	1,926	2,611
R-squared	0.03	0.03	0.04	0.04	0.03
Control mean	0.19	0.19	0.19	0.19	0.19
q value (treatment)	0.00	0.46	0.48	0.83	0.83
q value (interaction)	0.83	0.83	1.00	1.00	1.00
Panel C: Outcome = income [ESQ401a/BSQ503]					
treatment	-2.2 (1.7)	-0.4 (1.8)	0.8 (2.2)	-1.3 (1.8)	2.9 (3.6)
treatment x interaction	2.8 (5.0)	1.9 (5.2)	-1.9 (4.1)	2.0 (5.1)	-3.3 (3.7)
N	2,869	2,869	1,918	1,918	2,600
R-squared	0.07	0.07	0.09	0.09	0.07
Control mean	8.7	8.7	8.7	8.7	8.7
q value (treatment)	0.9	1.0	1.0	1.0	1.0
q value (interaction)	1.0	1.0	1.0	1.0	1.0

Table shows estimates of equation (4), including potential mechanism and main effect of term interacted with treatment. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. Income measured in USD (using exchange rate from July 1, 2018), previous two months. Proportion trainings/exchange visits attended measured at school level. Active instruction time (column 3) is the proportion of classroom time in Q&A/discussion, student presentation, and project/interactive activity. Active instructional techniques observed (column 4) include group discussion, research, case study, role play, debate, finance/practice activity. Interaction terms in columns 3-4 are from teacher endline survey, using matched student-teacher pairs only. Active instructional techniques from student reports is mean of indicators for at least one use of small group work, presentation to class, and small group presentation to class.

Table A1: Youth education and economic activity, 2012 Rwanda Census

Sample	<u>National</u>			<u>Sample Districts</u>			<u>Non-sample districts</u>			<u>Non-sample districts (excluding Kigali)</u>		
	<u>15-19</u>	<u>20-24</u>	<u>15-24</u>	<u>15-19</u>	<u>20-24</u>	<u>15-24</u>	<u>15-19</u>	<u>20-24</u>	<u>15-24</u>	<u>15-19</u>	<u>20-24</u>	<u>15-24</u>
Age group												
<u>Education</u>												
never attended school	6%	11%	8%	6%	12%	9%	5%	11%	8%	6%	12%	9%
attended primary school	62%	55%	59%	64%	59%	61%	61%	53%	57%	64%	56%	60%
attended secondary school	32%	29%	31%	30%	27%	28%	33%	31%	32%	30%	29%	29%
attended university	0.2%	3.8%	1.9%	0.1%	2.1%	1.0%	0.2%	4.7%	2.4%	0.1%	2.8%	1.3%
currently attending school	63%	24%	44%	61%	21%	42%	63%	26%	45%	64%	25%	46%
literacy	86%	82%	84%	85%	80%	83%	86%	82%	84%	85%	80%	82%
<u>Economic activity</u>												
employed	25%	54%	39%	27%	59%	42%	23%	52%	37%	23%	53%	37%
self-employed/family worker	73%	75%	74%	81%	83%	82%	68%	70%	69%	77%	79%	79%
agriculture sector	67%	67%	67%	76%	77%	77%	61%	60%	61%	73%	73%	73%

Data from 2012 Rwanda Census. Employed refers to those who answered "Yes" when asked "Aside from your own housework, did you work at least 1 hour during the last 7 days preceding the census night?" Self-employed and employed in agriculture condition on employment. Literacy refers to the ability to read and write with understanding at least one language.

Table A2: Student academic outcomes and entrepreneurship skills by gender (Table 5)

	Academic skills				Entrepreneurship skills				
	exam scores		monthly discount rate		understands	savings		entrepreneurship	
	overall	S6 entrepreneurship	S4/S5 promotional	<100%	<300%	compound interest	any amount (if >0)	knowledge	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: male students									
treatment	N/A	-0.10 (0.06)*	-0.07 (0.07)	0.03 (0.02)	0.07 (0.03)**	-0.03 (0.03)	-0.02 (0.03)	-4.9 (3.1)	-0.01 (0.01)
N		1,052	1,159	1,225	1,223	1,226	1,226	795	1,226
R-squared		0.42	0.26	0.03	0.04	0.02	0.06	0.10	0.05
Control mean		0.19	0.14	0.12	0.30	0.80	0.66	34.0	0.72
Baseline mean		-0.29	-0.26	0.20	0.39	0.65	0.38		0.68
q-value		1.00	1.00	1.00	0.65	1.00	1.00		1.00
Lee bounds		[-0.15,-0.01]	[-0.14,0.02]	[-0.01,0.03]	[-0.02,0.03]	[-0.03,0.01]	[-0.03,0.02]	[-3.1,10.2]	[-0.02,0.01]
Panel B: female students									
treatment	N/A	0.08 (0.07)	0.02 (0.06)	-0.01 (0.02)	0.02 (0.03)	0.01 (0.02)	0.05 (0.03)*	-4.3 (2.0)**	0.01 (0.01)
N		1,309	1,533	1,654	1,654	1,654	1,654	1,033	1,654
R-squared		0.47	0.28	0.04	0.05	0.02	0.08	0.11	0.05
Control mean		-0.19	-0.13	0.16	0.34	0.81	0.60	21.3	0.66
Baseline mean		-0.01	0.07	0.27	0.42	0.66	0.27	0.3	0.64
q-value		1.00	1.00	1.00	1.00	1.00	0.83	0.7	1.00
Lee bounds		[0.03,0.12]	[-0.02,0.06]	[-0.02,0.01]	[-0.01,0.01]	[-0.013,0.015]	[-0.01,0.01]	[-3.5,1.3]	[-0.01,0.009]

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. q-value reports sharpened q-value (Benjamini et al 2006). * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Exam scores normalized by control group mean and standard deviation. Baseline outcome for academic outcomes (columns 1-3) is S3 exam score. Baseline outcome for column (8) is dummies for savings between RWF1-4,999, RWF5,000-10,000, or more than RWF10,000 (not reported in table because this is a vector). Monthly discount rate based on stated preference for 5,000RWF today versus larger amount one month from now. Savings in column (8) reported in USD using exchange rate on July 1, 2018, winsorized at 99th percentile. Entrepreneurship knowledge is proportion correct on 7-item test.

Table A3: Student non-cognitive skills, by gender

	<u>university or beyond</u> (1)	<u>Aspirations business or professional</u> (2)	<u>business creation</u> (3)	<u>Locus of control</u> (4)	<u>Grit</u> (5)
<u>Panel A: male students</u>					
treatment	0.01 (0.02)	0.00 (0.03)	-0.01 (0.01)	-0.02 (0.08)	-0.03 (0.05)
N	1,226	1,226	1,226	1,226	1,226
R-squared	0.06	0.04	0.01	0.04	0.05
Control mean	0.82	0.66	0.98	6.50	3.37
Baseline mean	0.71	0.56	0.80	5.46	2.93
q-value	1.00	1.00	1.00	1.00	1.00
Lee bounds	[-0.03,0.02]	[-0.03,0.02]	[-0.03,-0.007]	[-0.14,0.09]	[-0.09,0.03]
<u>Panel B: female students</u>					
treatment	0.01 (0.03)	0.02 (0.03)	-0.01 (0.01)	-0.01 (0.08)	-0.03 (0.04)
N	1,654	1,654	1,654	1,654	1,654
R-squared	0.06	0.03	0.03	0.05	0.07
Control mean	0.77	0.65	0.98	6.29	3.38
Baseline mean	0.74	0.62	0.75	5.24	2.95
q-value	1.00	1.00	1.00	1.00	1.00
Lee bounds	[0.004,0.03]	[0.016,0.036]	[-0.011,0.008]	[-0.06,0.05]	[-0.06,0.01]

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. q-value reports sharpened q-value (Benjamini et al 2006). * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Locus of control is mean of 5 items about personal control over outcomes, with each item scaled from 1 (no control) to 10 (total control). Grit is mean of 4 items about personal persistence, scaled from 1 (not true) to 5 (very true).

Table A4: Student economic activity by gender

	<u>dropped</u> <u>out</u>	<u>all</u>	<u>business participation</u>		<u>family/</u> <u>peers</u>	<u>business characteristics</u>		<u>employment</u>	<u>income</u>	
			<u>own</u>	<u>student</u> <u>club</u>		<u>non-</u> <u>agricultural</u>	<u>has paid</u> <u>employees</u>		<u>total</u>	<u>business</u> <u>profit</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<u>Panel A: male students</u>										
treatment	0.00 (0.02)	0.06 (0.03)*	0.05 (0.03)	0.03 (0.01)***	0.00 (0.01)	-0.03 (0.02)*	0.01 (0.02)	-0.04 (0.03)	0.5 (1.4)	-20.0 (7.9)**
N	1,226	1,226	1,226	1,226	1,226	1,226	1,226	1,226	1,225	404
R-squared	0.02	0.07	0.07	0.02	0.01	0.04	0.03	0.04	0.11	0.10
Control mean	0.06	0.34	0.30	0.01	0.05	0.11	0.09	0.25	12.5	43.7
Baseline mean	0.00	0.29	0.29	0.00	0.11	0.06	0.29	0.19	2.5	6.6
q-value	0.78	0.26	0.31	0.08	0.78	0.26	0.57	0.31	0.7	0.1
Lee bounds	[-0.01,0.04]	[0.04,0.09]	[0.03,0.08]	[0.02,0.04]	[-0.01,0.03]	[-0.04,0.003]	[0.002,0.05]	[-0.06,-0.01]	[-0.49,3.2]	[-21.4,-3.4]
<u>Panel B: female students</u>										
treatment	0.00 (0.02)	0.04 (0.03)	0.02 (0.03)	0.04 (0.01)***	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	-0.04 (0.02)**	-0.7 (0.6)	-12.0 (8.2)
N	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,644	385
R-squared	0.03	0.06	0.06	0.03	0.02	0.04	0.03	0.04	0.02	0.05
Control mean	0.12	0.26	0.23	0.01	0.04	0.08	0.05	0.15	5.2	31.1
Baseline mean	0.00	0.17	0.17	0.00	0.08	0.03	0.17	0.09	0.8	2.8
q-value	0.78	0.31	0.57	0.00	0.57	0.41	0.59	0.17	0.4	0.3
Lee bounds	[-0.02,0.000]	[0.03,0.05]	[0.001,0.024]	[0.02,0.033]	[-0.01,0.01]	[-0.004,0.02]	[-0.01,0.01]	[-0.05,-0.03]	[-1.8,-0.5]	[-22.3,-6.5]

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Business involvement (all) is indicator for choice (a) or (c) in item ESQ401. Business characteristics refer to main business only. Baseline outcome for "business has employees" is indicator for involvement in any business. Employment is indicator for choice (b) or © in item BSQ401. Income measured in USD (using exchange rate from July 1, 2018), previous two months. Total income (column 9) uses self-reported total. Business profit (column 10) conditions on business involvement. Business profit adjusted by reported frequency of business earnings to estimate profits over two months. Alternative measure of total income is sum of reported employment income and business, adjusted by reported frequencies to estimate income for two months. All financial variables winsorized at 99th percentile.

Table A5: Student academic outcomes and entrepreneurship skills, controlling for baseline imbalance

	<u>Academic skills</u>			<u>Entrepreneurship skills</u>					
	<u>exam scores</u>			<u>monthly</u>		<u>understands</u>	<u>savings</u>	<u>entrepreneurship</u>	
	<u>overall</u>	<u>S6</u>	<u>S4/S5</u>	<u>discount rate</u>	<u><100%</u>	<u><300%</u>	<u>compound</u>	<u>any</u>	<u>amount</u>
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(9)
treatment	N/A	-0.01 (0.05)	-0.03 (0.06)	0.01 (0.02)	0.04 (0.02)**	-0.01 (0.02)	0.02 (0.02)	-4.4 (1.6)***	0.00 (0.01)
N		2,361	2,692	2,879	2,877	2,880	2,880	1,828	2,880
R-squared		0.45	0.27	0.03	0.04	0.02	0.06	0.12	0.05
Control mean		0.00	0.00	0.14	0.32	0.81	0.63	27.6	0.69
Baseline mean		-0.13	-0.07	0.24	0.40	0.65	0.32	N/A	0.66
q-value		1.00	1.00	1.00	0.17	1.00	1.00	0.12	1.00
Lee bounds		[-0.02,-0.007]	[-0.03,-0.02]	[-0.001,0.001]	[-0.001,0.001]	[-0.001,0.001]	[-0.001,0.001]	[-0.04,0.11]	[-0.001,0.0003]

Table shows estimates of equation (4), including baseline outcome where indicated. All regressions include controls for variables imbalanced at baseline (female, employed during school holiday, grit). Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Exam scores normalized to control mean and standard deviation. Baseline outcome for academic outcomes (columns 1-3) is S3 exam score. Baseline outcome for column (8) is dummies for savings between RWF1-4,999, RWF5,000-10,000, or more than RWF10,000 (not reported in table because this is a vector). Monthly discount rate based on stated preference for 5,000RWF today versus larger amount one month from now. Savings in column (8) reported in USD using exchange rate on July 1, 2018, winsorized at 99th percentile. Entrepreneurship knowledge is proportion correct on 7-item test. Results using administrative data will be omitted if data unavailable.

Table A6: Student non-cognitive skills, controlling for baseline imbalances

	<u>university or beyond</u>	<u>Aspirations business or professional</u>	<u>business creation</u>	<u>Locus of control</u>	<u>Grit</u>
	(1)	(2)	(3)	(4)	(5)
treatment	0.01 (0.02)	0.01 (0.02)	-0.01 (0.01)	-0.02 (0.06)	-0.03 (0.04)
N	2,880	2,880	2,880	2,880	2,880
R-squared	0.06	0.03	0.02	0.04	0.05
Control mean	0.80	0.66	0.98	6.39	3.37
Baseline mean	0.73	0.60	0.77	5.33	2.94
q-value	1.00	1.00	0.90	1.00	1.00
Lee bounds	[0.006,0.008]	[0.012,0.014]	[-0.01,-0.01]	[-0.022,-0.01]	[-0.03,-0.02]

Table shows estimates of equation (4), including baseline outcome where indicated. All regressions include controls for variables imbalanced at baseline (female, employed during school holiday, grit). Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Locus of control is mean of 5 items about personal control over outcomes, with each item scaled from 1 (no control) to 10 (total control). Grit is mean of 4 items about personal persistence, scaled from 1 (not true) to 5 (very true).

Table A7: Student economic activity, controlling for baseline imbalances

	<u>dropped</u> <u>out</u>	<u>all</u>	<u>business participation</u>		<u>family/</u> <u>peers</u>	<u>business characteristics</u>		<u>employment</u>	<u>income</u>	
			<u>own</u>	<u>student</u> <u>club</u>		<u>non-</u> <u>agricultural</u>	<u>has paid</u> <u>employees</u>		<u>total</u>	<u>business</u> <u>profit</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
treatment	0.00 (0.01)	0.05 (0.02)**	0.04 (0.02)	0.03 (0.01)***	0.01 (0.01)	-0.01 (0.01)	0.01 (0.01)	-0.04 (0.02)**	-0.1 (0.7)	-15.2 (5.9)**
N	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,869	789
R-squared	0.03	0.07	0.07	0.02	0.01	0.03	0.03	0.05	0.11	0.05
Control mean	0.09	0.30	0.27	0.01	0.04	0.09	0.07	0.19	8.7	38.2
Baseline mean	0.00	0.22	0.22	0.00	0.09	0.04	0.22	0.13	1.5	4.7
q-value	0.77	0.05	0.19	0.00	0.50	0.59	0.40	0.08	0.74	0.05
Lee bounds	[0.001,0.002]	[0.05,0.052]	[0.033,0.035]	[0.028,0.03]	[0.004,0.006]	[-0.005,-0.003]	[0.009,0.01]	[-0.036,-0.034]	[-0.2,0.02]	[-20.1,-12.8]

Table shows estimates of equation (4), including baseline outcome where indicated. All regressions include controls for variables imbalanced at baseline (female, employed during school holiday, grit). Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Business involvement (all) is indicator for choice (a) or (c) in item ESQ401. Business characteristics refer to main business only. Baseline outcome for "business has employees" is indicator for involvement in any business. Employment is indicator for choice (b) or © in item BSQ401. Income measured in USD (using exchange rate from July 1, 2018), previous two months. Total income (column 9) uses self-reported total. Business profit (column 10) conditions on business involvement. Business profit adjusted by reported frequency of business earnings to estimate profits over two months. Alternative measure of total income is sum of reported employment income and business, adjusted by reported frequencies to estimate income for two months. All financial variables winsorized at 99th percentile.

Table A8: Program take-up and curricular implementation – IV estimates counterpart of Table 3

	<u>teacher curricular implementation</u>			<u>administrator perceptions</u>	
	<u>Skills Lab scheduled</u>	<u>lesson plans</u>	<u>entrepreneurship knowledge</u>	<u>student-centered teaching</u>	<u>skill-based learning</u>
	(1)	(2)	(3)	(4)	(5)
exchange visit attendance	0.69 (0.07)***	-0.05 (0.06)	0.10 (0.03)***	0.19 (0.17)	0.03 (0.04)
N	235	333	333	204	204
1st stage F-stat	646.6	824.9	830.8	992.5	1081.3
Control mean	0.08	0.61	0.58	3.63	0.93
Baseline mean	0.10	0.59	0.82	3.25	0.54
q-value	0.00	0.32	0.00	0.32	0.32

Table shows estimates of equation (6), including baseline outcome where indicated. Exchange visit attendance (proportion) instrumented by treatment assignment. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Training and exchange visit attendance (columns 1-2) measure proportion of teacher training or exchange visits attended by any teacher from school (of 6 total), 2016-2018. Skills Lab scheduled (column 3) is an indicator for whether the teacher reports Skills Lab in its weekly schedule when enumerators called to plan their endline school visit. Baseline outcome for Skills Lab scheduled is knowledge of Skills Lab definition. Lesson plans (column 4) is mean of indicators for showing entrepreneurship lesson plan and notes. Entrepreneurship knowledge (column 5) is mean of questions about entrepreneurship curriculum. Column 6 is index of head teacher support for student-centered teaching methods (0=least supporting, 6=most supportive). Student-centered pedagogical tools include question and answer; group work; games; activities outside classroom; experiment; portfolio. Column 7 is indicator for whether head teacher agrees that teachers' primary goal should be to develop student skills. Baseline outcome for column 7 is mean of questions on support for skill-based learning (BHQ606/608/610).

Table A9: Pedagogical change – IV estimates counterpart of Table 4

	active instruction time						active instructional techniques			
	all observed classes			Skills Labs only			observed		student reports	
	full	first	second	full	first	second	all	Skills Labs	narrow	broad
	observation	half	half	observation	half	half	classes	only		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
exchange visit attendance	0.02	-0.09	0.13	-0.35	-0.14	-0.61	0.08	-0.06	0.03	0.09
	(0.04)	(0.04)**	(0.05)**	(0.19)*	(0.15)	(0.31)**	(0.03)***	(0.17)	(0.01)***	(0.01)***
N	235	235	235	70	70	70	235	70	2,611	2,611
1st stage F-stat	599.9	599.9	599.9	9.0	9.0	9.0	644.2	9.5	1083.9	1083.9
Control mean	0.48	0.45	0.52	0.64	0.52	0.78	0.31	0.31	0.95	0.62
q-value	0.28	0.05	0.03	0.06	0.17	0.05	0.02	0.28	0.00	0.00

Table shows estimates of equation (6), including baseline outcome where indicated. Exchange visit attendance (proportion, by school) instrumented by treatment assignment. Standard errors in parenthesis (robust for columns 1-8, clustered by school for columns 9-10). * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Active instruction time is the proportion of classroom time in Q&A/discussion, student presentation, and project/interactive activity. The baseline outcome for columns (1)-(6) is an index of support for interactive teaching. Defined as index = BTQ300 + (3-BTQ301) + BTQ302. BTQ301 enters negatively because it is number of interactive techniques among least comfortable for teacher. Scale of raw index is 0-12 (least to most interactive). Active instructional techniques observed (columns 7-8) include group discussion, research, case study, role play, debate, finance/practice activity. Active instructional techniques from student reports: "narrow" measure in column 8 is mean of indicators for at least one use of small group work, presentation to class, and small group presentation to class. "Broad" measure in column 10 is mean of all indicators from narrow measure, plus outside class activity, group discussion, research, case study, role play, financial calculations, computer work, and career planning.

Table A10: Student academic and entrepreneurial skills – IV estimates counterpart of Table 5

	Academic skills			Entrepreneurship skills					
	exam scores			monthly		understands	savings		entrepreneurship
	overall	S6	S4/S5	discount rate	discount rate	compound	any	amount	knowledge
(1)	entrepreneurship	promotional	<100%	<300%	interest	(7)	(if >0)	(9)	
exchange visit attendance	N/A	-0.01 (0.07)	-0.04 (0.07)	0.02 (0.02)	0.06 (0.02)**	-0.01 (0.02)	0.03 (0.03)	-6.8 (2.0)***	0.00 (0.01)
N		2,361	2,692	2,879	2,877	2,880	2,880	1,828	2,880
1st stage F-stat		1,008.6	1,021.5	1,048.8	1,044.4	1,044.6	1,054.8	1,079.2	1,043.5
Control mean		0.00	0.00	0.14	0.32	0.81	0.63	27.6	0.69
Baseline mean		-0.13	-0.07	0.24	0.40	0.65	0.32	0.32	0.66
q-value		1.00	1.00	1.00	0.09	1.00	1.00	0.01	1.00

Table shows estimates of equation (6), including baseline outcome where indicated. Exchange visit attendance (proportion, by school) instrumented by treatment assignment. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Exam scores normalized to control mean and standard deviation. Baseline outcome for academic outcomes (columns 1-3) is S3 exam score. Baseline outcome for column (8) is dummies for savings between RWF1-4,999, RWF5,000-10,000, or more than RWF10,000 (not reported in table because this is a vector). Monthly discount rate based on stated preference for 5,000RWF today versus larger amount one month from now. Savings in column (8) reported in USD using exchange rate on July 1, 2018, winsorized at 99th percentile. Entrepreneurship knowledge is proportion correct on 7-item test. Results using administrative data will be omitted if data unavailable.

Table A11: Student non-cognitive skills – IV estimates counterpart of Table 6

	<u>university or beyond</u>	<u>Aspirations business or professional</u>	<u>business creation</u>	<u>Locus of control</u>	<u>Grit</u>
	(1)	(2)	(3)	(4)	(5)
exchange visit attendance	0.01 (0.02)	0.02 (0.03)	-0.01 (0.01)	-0.04 (0.08)	-0.03 (0.05)
N	2,880	2,880	2,880	2,880	2,880
1st stage F-stat	1,055.6	1,054.0	1,053.9	1,045.8	1,022.6
Control mean	0.80	0.66	0.98	6.39	3.37
Baseline mean	0.73	0.60	0.77	5.33	2.94
q-value	1.00	1.00	1.00	1.00	1.00

Table shows estimates of equation (6), including baseline outcome where indicated. Exchange visit attendance (proportion, by school) instrumented by treatment assignment. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Lee bounds calculated by regressing outcome on baseline outcome and randomization strata dummies, then calculating bounds for residuals. Locus of control is mean of 5 items about personal control over outcomes, with each item scaled from 1 (no control) to 10 (total control). Grit is mean of 4 items about personal persistence, scaled from 1 (not true) to 5 (very true).

Table A12: Student economic activity – IV estimates counterpart of Table 7

	<u>dropped</u>		<u>business participation</u>			<u>business characteristics</u>		<u>employment</u>	<u>income</u>	
	<u>out</u>	<u>all</u>	<u>own</u>	<u>student</u>	<u>family/</u>	<u>non-</u>	<u>has paid</u>		<u>total</u>	<u>business</u>
	(1)	(2)	(3)	<u>club</u>	<u>peers</u>	<u>agricultural</u>	<u>employees</u>	(8)	(9)	<u>profit</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
exchange visit attendance	0.00 (0.02)	0.06 (0.03)**	0.03 (0.03)	0.04 (0.01)***	0.00 (0.01)	-0.01 (0.01)	0.01 (0.01)	-0.06 (0.02)***	-1.0 (0.9)	-19.9 (7.7)***
N	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,869	789
1st stage F-stat	1045.9	1066.1	1066.1	1046.5	1046.0	1052.8	1066.1	1073.2	1100.7	581.6
Control mean	0.09	0.30	0.27	0.01	0.04	0.09	0.07	0.19	8.7	38.2
Baseline mean	0.00	0.22	0.22	0.00	0.09	0.04	0.22	0.13	1.5	4.7
q-value	0.67	0.09	0.28	0.00	0.62	0.62	0.62	0.02	0.34	0.03

Table shows estimates of equation (6), including baseline outcome where indicated. Exchange visit attendance (proportion, by school) instrumented by treatment assignment. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Business involvement (all) is indicator for choice (a) or (c) in item ESQ401. Business characteristics refer to main business only. Baseline outcome for "business has employees" is indicator for involvement in any business. Employment is indicator for choice (b) or © in item BSQ401. Income measured in USD (using exchange rate from July 1, 2018), previous two months. Total income (column 9) uses self-reported total. Business profit (column 10) conditions on business involvement. Business profit adjusted by reported frequency of business earnings to estimate profits over two months. Alternative measure of total income is sum of reported employment income and business, adjusted by reported frequencies to estimate income for two months. All financial variables winsorized at 99th percentile.

Table A13: Heterogeneous treatment effects – IV estimates counterpart of Table 8

	Interaction term					
	<u>female</u> <u>student</u>	<u>baseline</u> <u>exam</u> <u>score</u>	<u>above</u> <u>median</u> <u>SES</u>	<u>female</u> <u>teacher</u>	<u>teacher</u> <u>experience</u>	<u>qualified</u> <u>teacher</u>
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Outcome = business involvement [ESQ401/BSQ600]						
exchange visit attendance	0.07 (0.04)*	0.06 (0.03)**	0.08 (0.04)**	0.06 (0.03)*	0.02 (0.04)	0.13 (0.05)***
exchange visits x interaction	-0.02 (0.05)	0.03 (0.02)	-0.04 (0.04)	-0.07 (0.06)	0.00 (0.00)	-0.15 (0.06)**
N	2,880	2,779	2,880	2,487	2,487	2,487
Control mean	0.30	0.30	0.30	0.30	0.30	0.30
q-value (treatment)	0.22	0.11	0.12	0.26	0.69	0.05
q-value (interaction)	0.69	0.40	0.56	0.40	0.75	0.10
Panel B: Outcome = employment [ESQ401/BSQ500]						
exchange visit attendance	-0.06 (0.04)	-0.06 (0.02)***	-0.04 (0.03)	-0.08 (0.03)***	-0.08 (0.03)***	-0.04 (0.04)
exchange visits x interaction	0.01 (0.04)	0.05 (0.02)***	-0.05 (0.04)	0.03 (0.05)	0.00 (0.00)	-0.04 (0.05)
N	2,880	2,779	2,880	2,487	2,487	2,487
Control mean	0.19	0.19	0.19	0.19	0.19	0.19
q-value (treatment)	0.35	0.05	0.36	0.05	0.05	0.40
q-value (interaction)	0.79	0.05	0.40	0.69	0.56	0.58
Panel C: Outcome = income [ESQ401a/BSQ503]						
exchange visit attendance	0.16 (1.69)	-0.93 (0.88)	0.34 (1.12)	-1.87 (1.23)	-2.14 (1.50)	-0.30 (1.41)
exchange visits x interaction	-0.74 (1.77)	0.07 (0.81)	-2.69 (1.45)*	1.76 (1.90)	0.13 (0.16)	-1.48 (1.83)
N	2,869	2,768	2,869	2,480	2,480	2,480
Control mean	8.7	8.6	8.7	8.5	8.5	8.5
q-value (treatment)	0.79	0.56	0.71	0.35	0.36	0.75
q-value (interaction)	0.69	0.79	0.22	0.56	0.58	0.58

Table shows estimates of equation (6), including baseline outcome and main effect of term interacted with treatment. Exchange visit attendance (proportion, by school) and its interaction instrumented by treatment assignment and its interaction. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Income measured in USD (using exchange rate from July 1, 2018), previous two months. All financial variables winsorized at 99th percentile. Baseline exam score normalized to mean zero and standard deviation one. SES is first principal component of household assets, parents' education, and indicator for parents in business or professional occupation. All interaction terms measured at baseline.

Table A14: Mechanisms – IV estimates counterpart of Table 9

	<u>Interaction term</u>		
	<u>active</u> <u>instruction</u> <u>time</u> (1)	<u>active</u> <u>instructional</u> <u>techniques</u> <u>(observed)</u> (2)	<u>active</u> <u>instructional</u> <u>techniques</u> <u>(student reports)</u> (3)
<u>Panel A: Outcome = business involvement [ESQ401/BSQ600]</u>			
exchange visit attendance	0.18 (0.11)	0.17 (0.09)*	0.23 (0.21)
exchange visits x interaction	-0.17 (0.23)	-0.20 (0.22)	-0.17 (0.22)
N	1,926	1,926	2,611
Control mean	0.30	0.30	0.28
q-value (treatment)	0.98	0.98	0.98
q-value (interaction)	0.98	0.98	0.98
<u>Panel B: Outcome = employment [ESQ401/BSQ500]</u>			
exchange visit attendance	-0.16 (0.09)*	-0.11 (0.07)	-0.24 (0.16)
exchange visits x interaction	0.16 (0.17)	0.05 (0.19)	0.18 (0.16)
N	1,926	1,926	2,611
Control mean	0.20	0.20	0.18
q-value (treatment)	0.98	0.98	0.98
q-value (interaction)	0.98	1.00	0.98
<u>Panel C: Outcome = income [ESQ401a/BSQ503]</u>			
exchange visit attendance	1.3 (3.1)	-1.7 (2.2)	4.2 (5.2)
exchange visits x interaction	-2.9 (6.0)	2.6 (6.1)	-4.7 (5.4)
N	1,918	1,918	2,600
Control mean	7.0	7.0	6.8
q-value (treatment)	1.00	0.98	0.98
q-value (interaction)	1.00	1.00	0.98

Table shows estimates of equation (6), including potential mechanism and main effect of term interacted with treatment. Exchange visit attendance (proportion, by school) and its interaction instrumented by treatment assignment and its interaction. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. q-value reports sharpened q-value (Benjamini et al 2006). Income measured in USD (using exchange rate from July 1, 2018), previous two months. Proportion trainings/exchange visits attended measured at school level. Active instruction time (column 3) is the proportion of classroom time in Q&A/discussion, student presentation, and project/interactive activity. Active instructional techniques observed (column 4) include group discussion, research, case study, role play, debate, finance/practice activity. Interaction terms in columns 3-4 are from teacher endline survey, using matched student-teacher pairs only. Active instructional techniques from student reports is mean of indicators for at least one use of small group work, presentation to class, and small group presentation to class.

Table A15: Pedagogical change, detail (hypothesis H2, exploratory)

	Outcome: proportion of teacher time spent in activity										
	demonstration/ lecture	Q&A/ discussion	student presentation	practice/ drill	project/ interactive activity	silent seat work	copying	instructions for assignments	uninvolved	classroom management	out of room
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: full observation											
treatment	0.01 (0.02)	0.02 (0.03)	0.00 (0.02)	-0.02 (0.01)	-0.01 (0.01)	0.01 (0.01)	-0.04 (0.01)***	-0.01 (0.02)	0.01 (0.01)*	0.01 (0.01)**	0.00 (0.00)
N	235	235	235	235	235	235	235	235	235	235	235
R-squared	0.12	0.10	0.13	0.08	0.11	0.09	0.13	0.07	0.13	0.13	0.07
Control mean	0.21	0.28	0.17	0.06	0.02	0.03	0.09	0.08	0.02	0.03	0.01
q-value	0.84	0.62	0.95	0.45	0.50	0.66	0.14	0.70	0.36	0.30	0.94
Panel B: first half of observation											
treatment	0.05 (0.03)*	-0.03 (0.03)	-0.02 (0.02)	0.00 (0.02)	-0.01 (0.01)	0.02 (0.01)	-0.03 (0.02)	0.01 (0.02)	0.00 (0.01)	0.01 (0.01)	0.00 (0.01)
N	235	235	235	235	235	235	235	235	235	235	235
R-squared	0.11	0.08	0.09	0.09	0.10	0.10	0.11	0.09	0.14	0.12	0.07
Control mean	0.25	0.32	0.10	0.07	0.01	0.02	0.09	0.08	0.01	0.03	0.01
q-value	0.36	0.50	0.66	0.95	0.66	0.36	0.36	0.86	0.66	0.44	0.95
Panel C: second half of observation											
treatment	-0.05 (0.03)*	0.10 (0.04)**	0.03 (0.04)	-0.03 (0.02)*	-0.01 (0.01)	0.00 (0.02)	-0.05 (0.02)***	-0.03 (0.02)	0.02 (0.01)	0.02 (0.01)*	0.00 (0.01)
N	235	235	235	235	235	235	235	235	235	235	235
R-squared	0.14	0.13	0.15	0.10	0.12	0.12	0.13	0.12	0.12	0.08	0.08
Control mean	0.17	0.25	0.25	0.04	0.02	0.04	0.09	0.08	0.03	0.04	0.01
q-value	0.36	0.17	0.66	0.36	0.44	0.95	0.14	0.44	0.36	0.36	0.95

Table shows estimates of equation (4), including baseline outcome where indicated. Robust standard errors in parenthesis. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. Outcome is the proportion of classroom time in listed activity, based on 52-minute observation. The baseline outcome is an index of support for interactive teaching. Defined as index = BTQ300 + (3-BTQ301) + BTQ302. BTQ301 enters negatively because it is number of interactive techniques among least comfortable for teacher. Scale of raw index is 0-12 (least to most interactive).

Table A16: Student economic activity, savings/profits/expenses (hypothesis H4, exploratory)

outcome units	savings		profits		expenses	
	level (USD) conditional	level (USD) unconditional	level (USD) conditional	level (USD) unconditional	level (USD) conditional	level (USD) unconditional
	(1)	(2)	(3)	(4)	(5)	(6)
treatment	-5.4 (1.6)***	-2.7 (1.2)**	-15.8 (6.1)**	-2.7 (1.7)	-7.8 (25.3)	1.8 (7.0)
N	1,828	2,880	789	2,770	776	2,757
R-squared	0.1	0.09	0.05	0.02	0.07	0.02
Control mean	27.6	17.4	38.2	10.4	83.5	22.3
Baseline mean	N/A	N/A	4.4	2.4	N/A	N/A
q-value	0.03	0.11	0.10	0.29	0.10	0.29

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. Conditional estimates for savings condition on any savings. Conditional estimates for profits and expenses condition on business involvement. All outcomes winsorized at 99th percentile.

Table A17: Student economic activity, detailed business activity (hypothesis H4, exploratory)

sample activity	unconditional					conditional on business participation				
	written ideas	create product	kept financial records	paid tax	registered	written ideas	create product	kept financial records	paid tax	registered
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
treatment	0.02 (0.01)*	0.01 (0.01)	0.02 (0.01)*	-0.01 (0.01)	-0.01 (0.00)	0.04 (0.03)	-0.02 (0.04)	0.03 (0.03)	-0.03 (0.02)*	-0.02 (0.01)*
N	2,880	2,880	2,880	2,880	2,880	899	899	899	899	899
R-squared	0.03	0.02	0.01	0.01	0.01	0.03	0.02	0.02	0.03	0.02
Control mean	0.09	0.10	0.05	0.02	0.01	0.29	0.32	0.18	0.07	0.04
q-value	0.22	0.53	0.22	0.39	0.35	0.35	0.53	0.48	0.22	0.25

Table shows estimates of equation (4). Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. All regressions include baseline business ownership as baseline outcome.

Table A18: Student economic activity, by quartile of likely business participation (hypothesis H4, exploratory)

	<u>dropped</u>	<u>business participation</u>				<u>business characteristics</u>		<u>employment</u>	<u>income</u>		<u>income</u>
	<u>out</u>	<u>all</u>	<u>own</u>	<u>student</u>	<u>family/</u>	<u>non-</u>	<u>has paid</u>		<u>total</u>	<u>business</u>	<u>(alternative</u>
	(1)	(2)	(3)	<u>club</u>	<u>peers</u>	<u>agricultural</u>	<u>employees</u>	(8)	(9)	<u>profit</u>	<u>measure)</u>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
treatment	0.01 (0.03)	0.05 (0.03)	0.01 (0.03)	0.02 (0.01)**	0.01 (0.01)	0.02 (0.02)	0.01 (0.01)	0.01 (0.02)	0.1 (0.8)	-3.7 (14.3)	0.2 (1.6)
treatment*quartile 2	0.02 (0.04)	0.06 (0.04)	0.08 (0.04)*	0.02 (0.02)	0.00 (0.02)	0.00 (0.03)	-0.01 (0.02)	-0.05 (0.03)	2.2 (1.5)	-19.2 (19.0)	-1.6 (2.5)
treatment*quartile 3	0.02 (0.04)	-0.03 (0.05)	-0.01 (0.05)	0.01 (0.01)	-0.01 (0.02)	-0.04 (0.03)	-0.01 (0.03)	-0.08 (0.04)*	-1.7 (1.7)	-7.5 (17.7)	-2.5 (3.1)
treatment*quartile 4	-0.06 (0.03)*	-0.02 (0.05)	0.01 (0.05)	0.00 (0.02)	0.00 (0.02)	-0.06 (0.03)*	0.00 (0.03)	-0.08 (0.04)*	-2.8 (1.7)	-15.3 (16.9)	-3.6 (2.8)
N	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,880	2,869	789	2,880
R-squared	0.02	0.11	0.11	0.02	0.01	0.04	0.03	0.05	0.10	0.05	0.05
Control mean	0.09	0.30	0.26	0.01	0.04	0.09	0.07	0.19	8.7	38.2	10.6
Baseline mean	0.00	0.22	0.22	0.00	0.09	0.04	0.22	0.13	1.5	4.7	1.5
q-value	1.00	1.00	1.00	0.33	1.00	1.00	1.00	1.00	1.00	1.00	1.00
q-value (Quartile 2)	0.60	0.55	0.12	0.15	1.00	1.00	1.00	0.15	0.15	0.55	0.60
q-value (Quartile 3)	1.00	1.00	1.00	1.00	1.00	0.48	1.00	0.13	0.48	1.00	0.82
q-value (Quartile 4)	0.12	1.00	1.00	1.00	1.00	0.12	1.00	0.12	0.12	0.40	0.28

Table shows estimates of equation (4), including baseline outcome where indicated. Standard errors in parentheses, clustered by school. * significant at 10%; ** significant at 5%; *** significant at 1%, based on conventional p-values. Quartiles based on propensity to start business at endline, calculated by logit regression of endline business participation on all school, teacher, and student baseline characteristics in pre-specified balance tests, baseline business ownership, and randomization strata. Quartile 4 is highest; quartile 1 is omitted category. Main effects also included in outcomes regression. Business involvement (all) is indicator for choice (a) or (c) in item ESQ401. Business characteristics refer to main business only. Baseline outcome for "business has employees" is indicator for involvement in any business. Employment is indicator for choice (b) or (c) in item BSQ401. Income measured in USD (using exchange rate from July 1, 2018), previous two months. Total income (column 9) uses self-reported total. Business profit (column 10) conditions on business involvement. Business profit adjusted by reported frequency of business earnings to estimate profits over two months. Alternative measure of total income is sum of reported employment income and business, adjusted by reported frequencies to estimate income for two months. All financial variables winsorized at 99th percentile.

Figures

Figure 1: Theory of Change

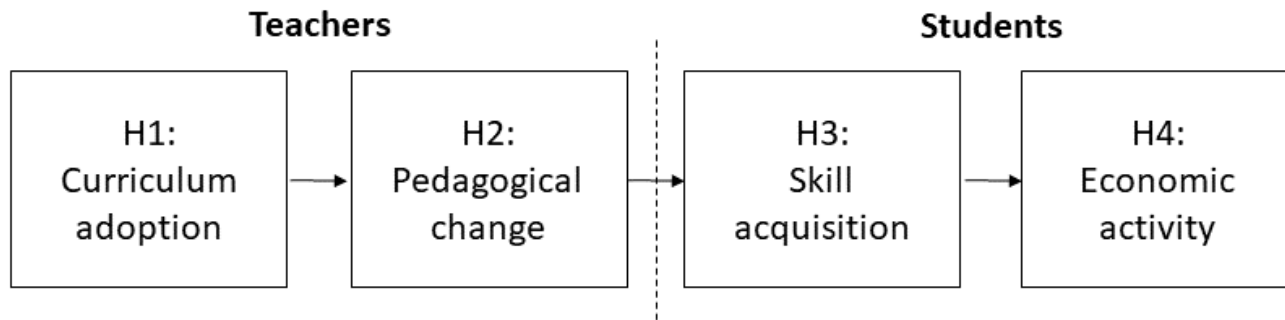


Figure 2: Schools in sample

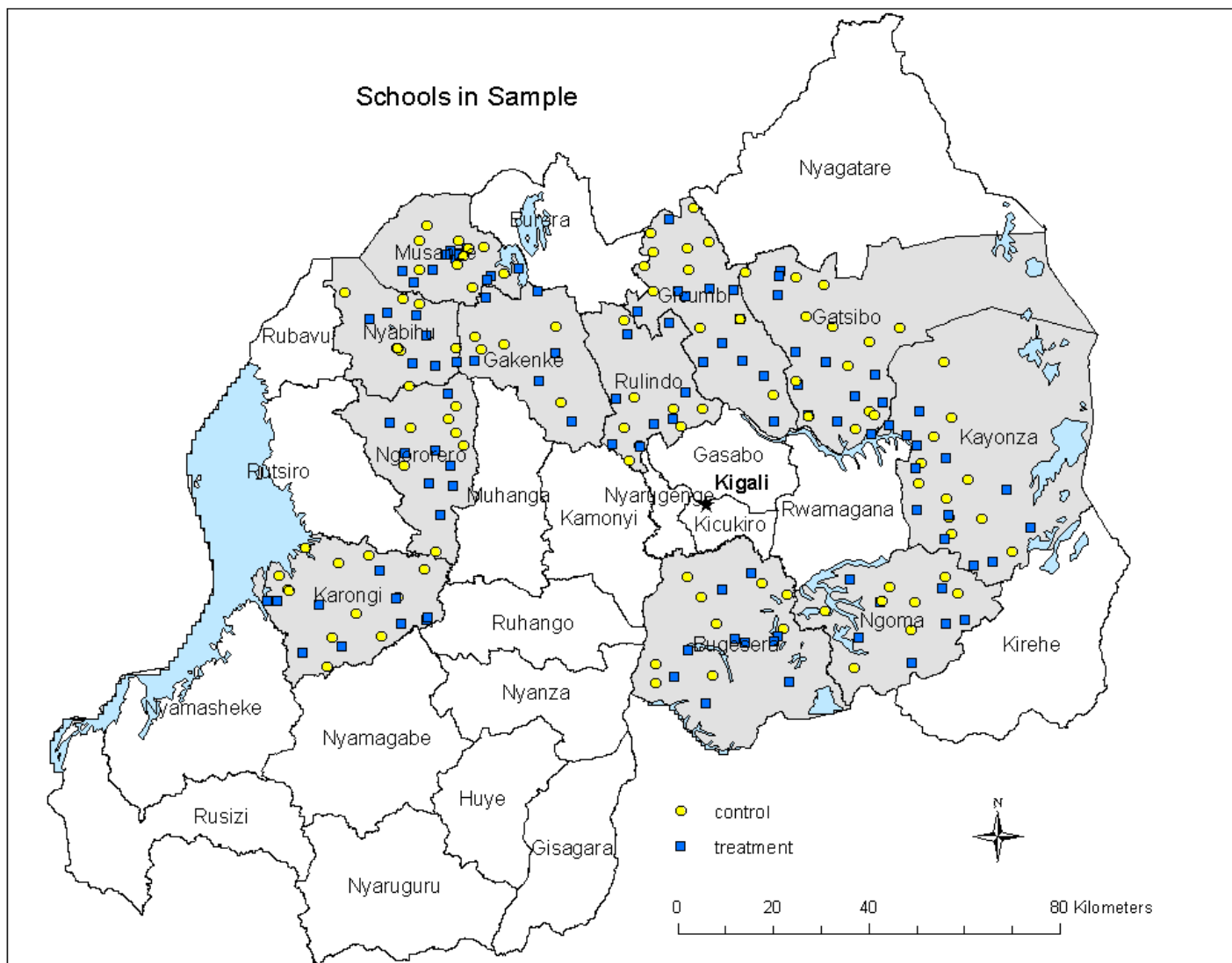


Figure 3: Timeline

