Journal of Development Economics 184 (2027) 103898

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

# Career guidance and investments in education: Experimental evidence from Cambodia

Esther Gehrke[a],*, Friederike Lenel[b], Claudia Schupp[c]

[a] Wageningen University and Research, Agricultural Economics and Rural Policy, Hollandseweg 1, Wageningen, 6706KN, The Netherlands
[b] Potsdam Institute for Climate Impact Research (PIK), Telegrafenberg A 31, Potsdam, 14473, Germany
[c] Technical University of Munich, Richard-Wagner-Strasse 1, Munich, 80333, Germany

ARTICLE INFO

Dataset link: 10.5281/zenodo.21027280

*JEL classification:*
C93
D83
D90
I21
O15

*Keywords:*
Aspirations
Career guidance
Education
Field experiment

ABSTRACT

Students in low-income contexts often lack guidance in their career decisions, which can lead to a misallocation of educational investments. We report on a randomized field experiment conducted with 1715 students in rural Cambodia and show that a half-day career guidance workshop designed to support adolescents in developing occupational aspirations on average increased educational investments. We document substantial heterogeneity in treatment effects by baseline student performance. While the workshop increased schooling efforts of high-performing students, treated low-performing students reduced their educational investments. We develop a simple model that explains this treatment effect heterogeneity by baseline performance.

## 1. Introduction

Access to education has improved substantially over the last decades in most low- and middle-income countries. However, a large proportion of students still drop out of school before completing compulsory education, and few continue with higher education (UNESCO, 2020). Education-related decisions are highly consequential for the types of careers children can pursue later in life and the living standards they can achieve. At the same time, these decisions are not easy; they need to be taken when children are relatively young and require substantial guidance (Heckman and Mosso, 2014). Without adequate guidance, students may not only have insufficient knowledge about the available educational pathways, but may also fail to develop aspirations that might enable them to pursue different educational and career paths than their parents, and escape the self-perpetuating cycles of poverty (Mani and Riley, 2021; van der Weide et al., 2024).

Personalized career guidance interventions for schoolchildren have the potential to shape aspirations, increase educational investments and thereby change schooling trajectories, as evidence from higher-income countries suggests (Renée, 2025; de Koning et al., 2025; Carlana et al., 2022). Successful career guidance interventions typically rely on a high degree of in-person contact, indicating that guidance is most effective when the information provided aligns with students' abilities and interests and clearly outlines the steps required to pursue a particular career. Thereby, goals become realistic and tangible and are more likely to translate into educational investments. As such interventions are costly and difficult to scale, they remain rare in low-income contexts, but are much needed (ILO, 2024; OECD, 2025): Parents often lack the knowledge and experience to properly support children in their educational and career related decisions, and teachers typically lack the incentives and resources to guide their students individually.

In this paper, we study the effectiveness of a half-day career guidance workshop targeting students of Grade 9, the final year of compulsory schooling, in rural Cambodia. The workshop was specifically designed to take into account students' individual interests and abilities, while remaining scalable. During the workshop, students first work through an interest and career exploration tool; an app that helps them reflect on their personal interests and allows them to explore (personalized) information about different careers that vary in their level of

* Corresponding author.
*E-mail address:* esther.gehrke@wur.nl (E. Gehrke).

[1] The information content of the careers is designed to introduce previously less known careers as well as to make possible career choices more tangible, for example by introducing the tasks that are typically involved in the daily activities, but also by stressing their societal relevance.

https://doi.org/10.1016/j.jdeveco.2026.103898
Received 12 January 2026; Received in revised form 30 June 2026; Accepted 3 August 2026
Available online 7 August 2026

required schooling.[1] Students then participate in an information session that presents paths both to higher education and to vocational training, and discusses academic requirements for attending high school, educational costs, and financing options. This low-cost and easily-scalable intervention aims to equip students with information about a range of occupations to match heterogeneous abilities and interests, while providing guidance on the next steps in their educational path, whether transitioning to high school or into vocational training.

Cambodia is a particularly relevant context for career guidance interventions. During the Khmer Rouge regime in the 1970s, the educational sector was systematically destroyed. Schools and universities were closed and educated people fled the country or were persecuted (UNESCO, 2011). With the subsequent Third Indochina War and a long period of internal conflict, the educational system was not reconstructed until the late 1990s. As a consequence, education levels among Cambodian adults are extremely low today, with severe repercussions on younger generations. Students lack information and guidance about career paths, and educational aspirations are often immensely unrealistic (Eng et al., 2014). Dropout rates are high, and the transition rate to high school is low, especially in rural areas (Ministry of Education, Youth and Sport, 2017).

We evaluate the effect of the intervention by exploiting the randomized assignment of 37 schools to either treatment or control status, with 1715 students participating in the study. The career guidance workshop was conducted in early 2020, shortly before schools were closed due to the COVID-19 pandemic. In terms of outcomes, we focus on schooling information from student-level administrative records collected throughout the last year of lower-secondary school and the subsequent three years of high school, as well as on self-reported information regarding study-behavior, and educational and occupational aspirations from a phone survey that was conducted about five months after the intervention and during the first COVID-19 lockdown.

We find that attending the career guidance workshop had positive effects on educational investments, which started to materialize about eight months after the intervention took place and persisted until high school completion. At the time of the phone survey, we find no statistically significant treatment effects on students' educational or occupational aspirations, nor on their confidence in reaching these aspirations. Students in treatment schools were also not more likely to study during the first lockdown period than their peers in control schools. However, treated students were 2.6 percentage points (pp) more likely to attend the final exam of Grade 9 (3% increase over the mean in the control group), performed better in the final exam conditional on participating (0.22 SD) and were approximately 5.9pp (7.9%) more likely to enroll in high school, 6.2pp (9.7%) more likely to progress to Grade 11 one year later, 5.6pp (9.2%) more likely to attend the midterm exam of Grade 11, and 3.9pp (7.0%) more likely to complete high school roughly 3.5 years after the intervention took place.

We then examine whether the intervention had differential impacts by students' ability. This is relevant as the workshop was explicitly designed to accommodate students' individual interests and abilities.[2] Using parametric and semi-parametric techniques, we examine treatment-effect heterogeneities along students' baseline academic performance, which we use as a proxy for students' abilities and which is by far the strongest predictor of subsequent schooling outcomes. We find that for low-performing students the intervention had negative effects in the short run (during Grade 9) on most outcomes and no effect in the medium run (during high school), while the intervention benefited high-performing students in both the short and the medium run. Treated students who performed in the bottom half of the grades distribution at baseline, studied less than their control group peers during the lockdown period, had (weakly) lower educational and occupational aspirations, but were no less confident in reaching their aspirations. Treated low-performers were also not less likely to attend the final exam of Grade 9. However, they performed worse in the exam conditional on attending. About one year after the intervention, the negative treatment effect on low-performers disappears, as few of them transitioned to high school in either treatment or control schools. By contrast, students who ranked in the upper half of the grades distribution at baseline were more likely to study during the lockdown period, had higher educational and occupational aspirations, and were (weakly) more confident in reaching their aspired education and occupation. They were more likely to attend the final exam of Grade 9 and performed better in that exam. High-performing treated students also kept outperforming their peers from control schools throughout high school and are driving the positive average treatment effects observed in the medium run.

The intervention was interrupted midway due to the COVID-19 pandemic, which triggered a lockdown and school closures in March 2020. This resulted in a smaller number of treated schools in our sample than what had been initially planned, and caused imbalances in baseline grades. The small number of clusters as well as the baseline imbalances are potential concerns to identification. To address these concerns, we perform a wealth of robustness checks. First, we show that average treatment effects and heterogeneities by baseline grades are robust to randomization inference, as well as to corrections for non-random attrition. Second, we trim outlier schools (below 10th and above 90th percentile) in average baseline grades from the sample. This procedure achieves balance in key variables, and if anything increases the significance of treatment effects and of the heterogeneities by grades. Finally, we show that the treatment effect heterogeneities are not driven by parental characteristics nor by school characteristics, and that the differential association between baseline grades and subsequent outcomes is non-existent before the intervention took place.

The documented heterogeneities in students' response to the career guidance workshop by baseline grades indicate that adjustments in students' educational investments coincide with adjustments in aspirations. One possible explanation is that the workshop allowed students to adjust their aspirations to levels that are better aligned with their academic performance. Participation in the workshop could have made low-performing students aware of alternative career paths that do not include higher education, thereby putting them in a position to adjust their educational (and occupational) aspirations to levels that are better aligned with their abilities and preferences. The decline in educational investments observed among these students during Grade 9 is consistent with them adjusting their educational investments to a level that is just sufficient to graduate lower-secondary school. By contrast, the workshop increased educational investments and academic performance among high-performing students compared to their control peers, which could be explained by an increase in aspirations also triggering more effort in this group of students.

To rationalize this mechanism, we develop a conceptual framework that defines aspirations as long-term goals, and in which students derive a milestone utility from achieving these goals. In particular, we combine insights from the models developed by Dalton et al. (2016) and Genicot and Ray (2017). Our framework features a rational agent who endogenously chooses an effort-aspiration pair that is aligned with their individual abilities and assumes that the agent formulates their *educational* aspirations to be consistent with their *occupational* aspirations. Aspirations (even though endogenously chosen) are drawn from an observed distribution of outcomes: the aspirations window. This window may be too narrow if students lack knowledge about career possibilities or misperceive the level of education necessary for a certain occupation. A truncated aspirations window can then trigger students to define educational aspirations that are not sufficiently aligned with their innate ability or preferences. Our model can explain why an intervention that provides students with information necessary

[2] For the purpose of this study, we understand student ability to capture cognition, as well as educational investments accumulated earlier in life.

to develop occupational aspirations that better align with their preferences and abilities affects educational aspirations and investments in such heterogeneous ways.

To the best of our knowledge, this is the first paper to evaluate a personalized career guidance intervention in a low-income context. Our paper speaks to three strands of literature. *First*, we contribute to the literature on information or mentoring interventions in educational contexts. Much of this literature focuses on high-income countries and evaluates the effect of providing educational and career guidance to students from disadvantaged backgrounds (Bettinger et al., 2012; Hoest et al., 2013; Hoxby and Turner, 2015; Goux et al., 2017; Abbiati et al., 2018; Kerr et al., 2020; Carlana et al., 2022; Renée, 2025). Studies that focus on low- and middle-income countries have shown that providing information about the returns to education can increase school attendance, improve test scores, and change educational trajectories of students (Nguyen, 2008; Jensen, 2010; Avitabile and de Hoyos, 2018).[3] The study most closely related to ours is Asri et al. (2025) who investigate the impact of a career exploration program targeting female students in urban high schools in India and find that while the program positively affected aspirations and planning in the short-term, it had no impact on educational trajectories. With our study, we show that personalized information on potential career paths and their educational requirements can indeed affect educational investments, indicating that the lack of information about career opportunities may lead students to sub-optimally invest in education.

*Second*, we contribute to the literature that investigates the role of students' aspirations in inducing educational investments. Most of this literature is based on role-model interventions (Bjorvatn et al., 2020; Bhan, 2020; Riley, 2024; Ahmed et al., 2024).[4] In contrast, our intervention encourages students to explore their personal interests and provides them with personalized information about a variety of potential career paths, making these careers more tangible. This approach supports the students in developing more diversified occupational and educational aspirations. Importantly, our insights may rationalize why role-model interventions are not always successful (see e.g. Kipchumba et al., 2023; Leight et al., 2021). Without being presented with a variety of educational paths and career possibilities, students may not be able to formulate new aspirations that are within their reach and may fail to adjust educational investments.

*Third*, we contribute to the theoretical literature that seeks to understand the reasons for aspiration failures (Dalton et al., 2016; Genicot and Ray, 2017, 2020; La Ferrara, 2019).[5] Our conceptual framework combines insights from Dalton et al. (2016) with those from Genicot and Ray (2017), and features a rational agent that chooses optimal effort-aspiration pairs but whose set of possible choices may be constrained due to information frictions. This model serves to highlight a new type of aspiration failure. If students lack information about career paths, their perceived set of possible effort-aspiration pairs is overly constrained, which leads to a misalignment between educational aspirations and ability, and to the misallocation of educational investments.[6]

The remainder of this paper proceeds as follows. In Section 2, we describe the setting and design of the intervention. The intervention's implementation and the data collection are described in Section 3. Section 4 presents the empirical approach and results. Section 5 discusses the underlying mechanisms and presents a simple conceptual framework that helps rationalize the evidence, and Section 6 concludes.

## 2. Setting

### 2.1. Education in Cambodia

The educational sector in Cambodia was systematically destroyed by the regime of the Khmer Rouge in the 1970s, during which the vast majority of teachers and academics fled the country or were killed (Chandler, 2007). The reconstruction of the educational sector did not start before the 1990s. The consequences are still visible today: most adults have not completed primary education (United Nations Development Programme, 2023); and while school completion rates have increased at primary and lower-secondary level, higher-secondary (high) school completion rates still lag behind (Huang et al., 2017).[7] The gross enrollment rate in lower-secondary schools is 56.5% and decreases to 28.1% in high schools. Furthermore, those students who manage to transition to high school are often not able to graduate with a diploma. During the school year of 2018–19, dropout rates in Grades 10, 11, and 12 were 14.1%, 7.2%, and 30.9%, respectively (Ministry of Education, Youth and Sport, 2019).

Such low enrollment can be partly explained by households' financial constraints. During Grade 9 most of the students reach their legal working age, 15, and can then support their family financially. Moreover, the expected costs of attending high school could deter students from considering further schooling. While general education is free, students have to cover the costs of material, extra classes, transportation, and accommodation in case the high school is located farther away. Our data suggest that there is a lot of uncertainty involved around these costs and that they are generally overestimated.

In addition, students seem to lack the necessary information and guidance from their parents and teachers to make informed educational decisions, which likely further explains the high dropout rates. Given their lack of education, parents can provide little support to their children in terms of homework or guidance in schooling decisions. Furthermore, parents often seem to underestimate the returns to education, most notably in low-income households (UNESCO, 2011). Overall, the involvement of parents or other family members in students' schooling is very rare (Benveniste et al., 2008). Teachers, on the other side, are not sufficiently compensated for providing personalized support and lack adequate training.

These insights are corroborated by findings from a pilot study in 2019, for which we surveyed students about their educational and occupational aspirations, their knowledge on career paths, and their beliefs about the costs associated with higher education.[8] Our findings suggest that students have little knowledge about potential career paths, and that this lack of information may constrain them in their educational decisions. Specifically, while all students were able to name a job they would like to do in the future, the range of different jobs mentioned is very limited. More than 85% of the students stated that

[3] Dizon-Ross (2019) studies an intervention that provides parents with information about their children's performance in school. She shows that the information improves the match between educational investments made by parents and their children's ability.

[4] For a recent review of various role-model interventions in education, see Serra (2025).

[5] The literature has so far identified two types of aspiration failures. *First*, Genicot and Ray (2017) consider a situation in which aspirations are exogenously drawn from a distribution of outcomes, the aspirations window. In such a setting, aspirations that are too low induce suboptimal effort, and aspirations that are too ambitious can lead to frustration because the goal becomes unachievable. *Second*, Dalton et al. (2016) consider a model in which aspirations are endogenously determined (and allowed to change over time), but individuals fail to internalize the feedback from effort to aspirations, and therefore choose suboptimally low aspirations.

[6] The importance of aligning aspirations to individual abilities has been highlighted *inter alia* in the context of sports (Lockwood and Kunda, 1997; Berger and Pope, 2011).

[7] The education system in Cambodia consists of six years of primary, three years of lower-secondary, and three years of high school. The first nine years of schooling are compulsory.

[8] We conducted surveys with 200 students and focus group discussions with 32 students who were at the end of Grade 8 and held interviews with teachers, parents, and local and international education experts.

they would like to become a teacher, general practitioner (doctor), police officer, or soldier.[9] At the same time, very few students demonstrated a clear understanding of how to reach their career goal, *i.e.*, what it requires in terms of schooling and where they would be able to pursue such studies. More than half of the students stated that they lack information about what they can do in the future. Talking to principals and teachers, it became apparent that future career options are not taught at school. Teachers admitted that they find it difficult to talk about career paths other than becoming a teacher as they have little knowledge about alternatives.

In other countries, there seems to be a similar concentration of occupational aspirations (Ross, 2019; Mann et al., 2020). Among adolescents who participated in PISA in 2018, for example, 53% (47%) of girls (boys) expect to work in one of the ten most commonly cited jobs. Even in neighboring Thailand, where educational attainment is substantially higher than in Cambodia, aspirations are quite concentrated, with 61% (64%) of girls (boys) mentioning one of the ten most commonly cited jobs (Mann et al., 2020).

### 2.2. *The career guidance workshop*

The intervention was designed as a half-day career guidance workshop tailored to Grade 9 students who are typically around 15 years old. The workshop caters to the diverse abilities and interests of the students. Rather than uniformly raising aspirations, which risks encouraging unrealistic goals, the workshop aims to help students identify career paths aligned with their interests and to make both familiar and less traditional options more tangible.

The workshop consists of three main parts, the first of which is an interest exploration tool (IET), which allows students to reflect on their interests and preferences, and reveals the students' congruence with different personality types. The second part consists of a career exploration tool (CET), in which students are provided with detailed information about a number of different jobs that they might find interesting. The third part is an information session on high schools and vocational training. For the first two parts, students work individually on a tablet (with the support of a research assistant if needed); the third part is conducted in-person in small groups.

For the IET, students work through three personality tests that are programmed in an electronic application. These tests are based on the theory of vocational interest developed by Holland (1959, 1997), and commonly known as the RIASEC model.[10] The theory of vocational interest is well established in psychology and management sciences, and posits that individuals who display personality traits that fit with their occupation display higher job satisfaction.[11] In our intervention, the personality tests serve the purpose of encouraging students to engage with their preferences, while allowing us to personalize the display of career options in the CET. The tests have been adapted to the Cambodian context by the project team in collaboration with local experts. During the tests, students are presented with statements on activities they might like or interests they might have (with pictures for illustration), and are asked to select the ones most applicable to them. The three tests differ in how statements are presented and how students can select them. After completion of all tests, the strongest personality types and a short description of what characterizes each type are revealed to each student based on their answers. It takes approximately 45 min to complete the tool. More details on the IET are presented in the Online Appendix B.1.

For the CET, students are shown a list of 18 occupations, three of which correspond to each personality type. For each personality type, the list contains one occupation that requires lower-secondary education plus some vocational training, one occupation that requires high school, and one occupation that requires a university degree. The list of jobs features occupations with which students are familiar, such as teacher, police officer, and general practitioner, as well as occupations that might not be known to the students but are relevant in the context, *i.e.*, agricultural technician, chef, or tour guide. The selection was guided by discussions with our local partners. We further ensured that all presented jobs existed in the local labor market and that there was sufficient demand for them by cross-checking that these occupations were reported by a sufficiently large number of workers in the Cambodia Socio-Economic Survey of 2017.[12]

The ordering of the occupations is personalized according to the students' strongest personality types. For each occupation, students can read more detail on the job content and main responsibilities, as well as the school subjects related to the occupation and the educational requirements.[13] Students can decide how much time they spend reading about each occupation.[14] For more details on the CET, see Online Appendix B.2.

The last part, the information session, provides detailed information on high schools and vocational training centers in the area, about the requirements and costs related to attending either institution, as well as scholarship possibilities. The content of this session is adapted to the context of each lower-secondary school, and conducted in person and interactively. For more details on the information session, see Online Appendix B.3. At the end of the intervention day, each student takes home a leaflet with all 18 occupations and their descriptions. Furthermore, teachers receive a poster on educational pathways that was discussed during the information session and are encouraged to place it on the wall of their class room.

## 3. Implementation and data

### 3.1. *Experimental design and timeline*

For the implementation of our intervention, we collaborated with Child's Dream (CD), an international NGO that offers high school scholarships in Northwest Cambodia. CD partners with 51 lower-secondary schools in 8 districts across 4 provinces (Battambang, Banteay Meanchey, Oddar Meanchey, and Siem Reap). The activities of CD in these schools are limited to handing out scholarship application materials to school principals to be distributed among interested students. For our study, we sampled all 39 schools that had a partnership with CD and a class size in Grade 9 above 30 students.[15] We expanded the sample by including 21 additional schools from the same provinces that are similar in characteristics to the CD partner schools. Of these 60

[9] These findings are replicated for the group of students targeted by our intervention, see Table A.1 (Figures A.1 to A.7 and Tables A.1 to A.21 are available in Online Appendix A).

[10] Holland (1997) proposes that there are six basic types of vocational interests, the "RIASEC" (Realistic, Investigative, Artistic, Social, Enterprising, and Conventional), that describe how people interact with their work environments and shape their career choices.

[11] There is ample empirical literature that confirms this. For recent reviews, see van Iddekinge et al. (2011) and Hoff et al. (2020).

[12] According to that same data, about 50% of workers in the country are wage employed. This share increases to 66% among workers below the age of 30.

[13] We did not include information on earnings prospects for two main reasons. First and foremost, we did not want students to anchor their reading choices on the earning potential, but rather consider the job content and how it matches their interests. Second, even if we would have wanted to include that information, reported earnings in Cambodia's Socio-Economic survey are scarce and exhibit large variance within occupation. This would have likely created false expectations.

[14] Students have a total of 17 min to read the descriptions that interest them, but can also log out earlier.

[15] In very few cases, the class size was below 30 but there was more than one class in Grade 9. In these cases, we combined two classes.

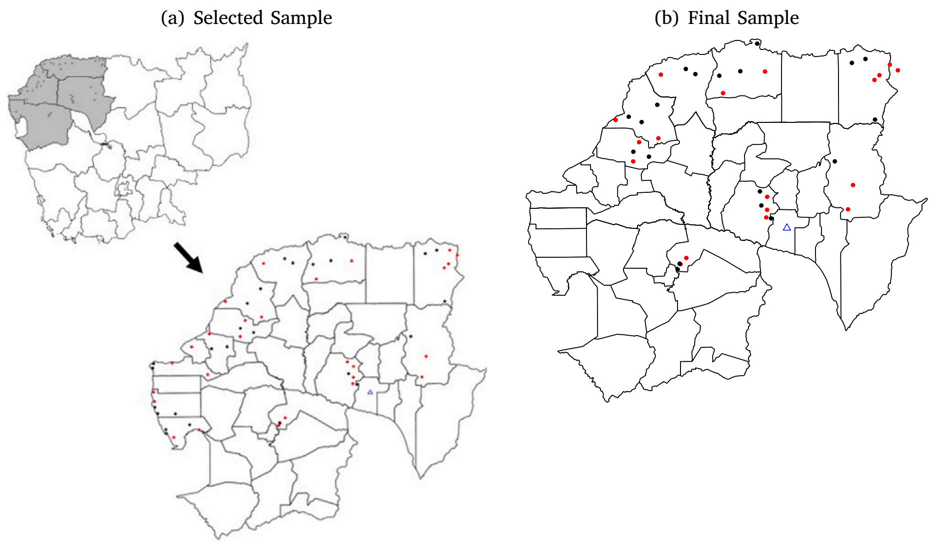


**Fig. 1.** Location of treatment and control schools.
*Notes*: Panel (a) shows the entire map of Cambodia in the upper left, highlighting the four provinces of interest in gray. The lower right map zooms into the four provinces, showing district borders and all initially selected treatment and control schools, marked in red and black respectively. Panel (b) highlights the location of the treatment and control schools in the final sample, again in red and black respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

schools, 30 were randomly assigned to receive the treatment (the half-day workshop); the remaining 30 served as control. For schools that had more than one class in Grade 9, we randomized the class that would receive the treatment in case of treatment schools (or serve as control in case of control schools). Randomization was stratified by district, but not by any school characteristics as this information was not yet available for all schools at the time. Fig. 1(a) depicts the location of the initial sample of treatment and control schools.[16]

The implementation of the intervention started in mid February 2020. By the beginning of April, the intervention was supposed to have been implemented in all 30 treatment schools, allowing some time between the intervention and the application deadlines for high school scholarships (including those awarded by CD). However, on March 16, 2020, as a measure to prevent the spread of COVID-19, the Cambodian government suddenly announced that all schools would be closed effective immediately. At that time, we had conducted the intervention in 18 schools in 8 districts. Our analysis therefore focuses on these 18 treatment schools and the 19 control schools that are located in the same districts (see Fig. 1(b) for the geographical location of these schools).[17] In the 18 treatment schools, 783 students out of the 862 invited students took part in the intervention.[18] Fig. 2 shows the timeline of the data collection and the sample composition. The collection of administrative data began in November 2019. At that time, there were in total 862 students in the treatment and 853 students in the control schools in the selected Grade 9 classes. We collected administrative data for all students, in particular gender, age, village of residence, as well as grades and absences for the months before the intervention was conducted. In addition, we collected teacher and school characteristics. On the day of the workshop, we also conducted three short surveys: one at the beginning (workshop baseline), one after students participated in the CET and before the information session (workshop midline), and one right after the information session (workshop endline).

In July to August 2020, we conducted a follow-up survey by phone. We reached 77% of the students (n= 1327). At that time, schools were still closed in Cambodia due to COVID-19.[19] In the phone survey, we asked students about their daily activities, their expectations, and aspirations in terms of education and future occupation. As a token of appreciation, students who finished the survey were eligible to participate in a lottery, and could choose between one of two potential prizes: phone credit or educational mentoring.

Schools reopened in September 2020 for ninth-graders, so that students could prepare for their final exam, which took place at the end of November 2020. The final exam grade determines whether students are allowed to enroll in high school. In December, after the final exams were graded, the government unexpectedly announced that all students who had registered for the final exam would obtain their Grade 9 diploma and would be allowed to transition to high

16 We randomly allocated students into one of three treatment arms in treatment schools: A1, A2, and A3 with respective chances 2:2:1 to study whether information take-up and processing differs when it is made self-relevant (see Badini et al., 2026). Students in A1 participated in all three parts of the workshop, students in A2 participated only in the career exploration tool (not personalized) and the information session, while students in A3 participated in the information session only. As such 80% of the treated students participated both in the CET as well as in the information session, which we consider to form the core of the intervention. More details on the within-school randomization are available in Online Appendix B.4.

17 We had organized the school visits to take place consecutively across the districts, i.e. all selected schools in one district were visited before moving to the next district. Given that randomization was stratified by district, and the timing of the school visits had been planned in early February, when COVID-19 was still unknown to most people, we can rule out any selection into treatment.

18 Students were informed about the workshop several days in advance, and they were told that they were free to participate or not. Students who did not show up on the day of the workshop display overall lower academic performance and more days of absence in the months before the intervention. During the workshop, a total of five students left before the end. In the remainder of this article, these five students are considered 'treated'; results do not change if these students are excluded.

19 Throughout the school closure, students were encouraged to study on their own with grade-specific TV programs. Furthermore, teachers were responsible for providing students with additional content and assignments. In a separate study, we show that learning activities varied greatly across students (Gehrke et al., 2023).

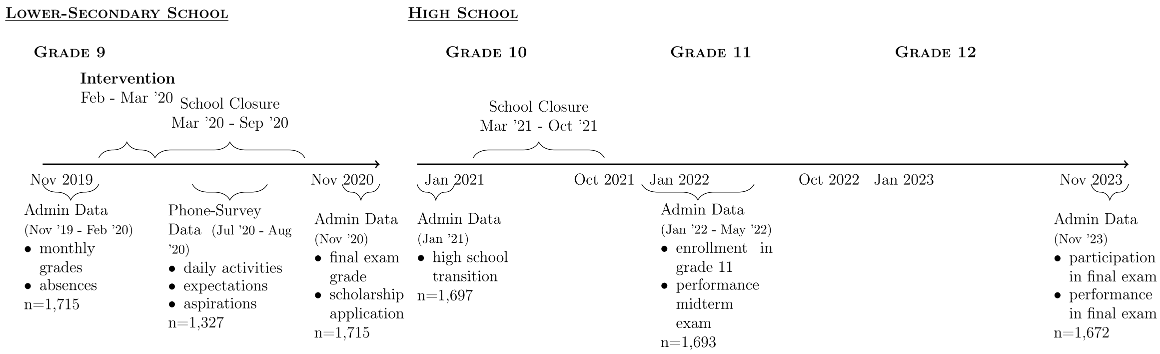


**Fig. 2.** Timeline of the data collection.

school, irrespective of their performance in the final exam. For that time period, we collected administrative information on participation in the final exam, which is our marker for whether a student dropped out during Grade 9, and final exam performance (actual grades given by the teacher before the government announcement). We also asked lower-secondary teachers if students requested the official transcripts necessary to enroll in high school. This gives us an important insight into who was intending to transition to high school (or to any other formal education). Furthermore, we collected data on scholarship applications from CD. This information is analyzed for those schools that partner with the NGO (28 out of 37 schools).

For tenth-graders, the new school year started in January 2021. However, all schools were closed again for a second lockdown at the end of February 2021 until the end of the school year in November 2021.[20] From the high schools, we collected data on whether students started Grade 10 (*i.e.*, transitioned to high school in January of 2021), whether they continued with Grade 11 in January of 2022 after the second lockdown, whether they attended and passed the midterm exam of Grade 11 (two years after our intervention) as well as data on high school completion in November of 2023. Importantly, high school completion is contingent on participation in the final exam, but not on the exam results.[21] The main selection thus seems to take place in the lead-up to the final exam. All in all, the data cover students' educational decisions more than three and a half years after the intervention.

### 3.2. Attrition

Of the 1715 students included in the study, 1317 were reached in the phone survey (77% of the original sample), and between 1672 and 1697 students (97%–99% of the original sample) were tracked with administrative data.[22] Participation in the phone survey was not random; in particular, female students and students who performed better in school at baseline are more likely to have participated in the phone survey. However, we do not find that attrition is systematically related to treatment status in either the survey or the administrative data (see Table A.2).

In the robustness checks, we address attrition in two ways. First, we compute Kling-Liebmann sensitivity bounds to account for missing values. Second, we test the sensitivity of the estimated effects on the survey-based outcomes by re-weighting the phone-survey sample with the inverse probability of participating in the phone survey.

### 3.3. Sample characteristics

The intervention targeted low-income students from rural areas. Table 1 presents student, school, and parental characteristics, as well as the balance between treatment and control groups in terms of these variables. This information is based on our two data sources: administrative data collected for all students in our sample before and after the intervention, and information from the phone survey that was conducted with students in treatment and control schools in summer 2020.

A little over half of the students in our sample are female, and they were on average 15 years old at the time of the intervention. Nearly all students lived in rural areas, on average 11 kilometers away from the district town. To reach their lower-secondary school, students needed to travel on average 3.6 kilometers and about ten kilometers to reach the next high school. In the phone survey, we asked students explicitly about their parents' education and occupation before the COVID-19 pandemic. From those students who knew their parents' education level, 81% (92%) reported that their father (mother) had completed primary education or less. At least one parent is a farmer for 69% of the students.

Baseline characteristics are overall balanced between treatment and control schools. However, out of 19 variables for which we have baseline values, 3 display differences in means that are statistically significant at the 10% level. Specifically, students in treatment schools live farther away from their school, have on average lower grades —measured as the sum of Math, English, and Khmer (the official language of Cambodia), averaged over the months December and January, and standardized across schools— and are being taught by younger teachers. The imbalance in the grade is the most worrisome as past grades are strong predictors of future educational outcomes. As there is substantial overlap in the grade distribution between treatment and

[20] During this second lockdown, classes were held online; however, attendance was not tracked.

[21] Those students who fail that exam (in our sample approximately 15% of the students) need to take extra classes if they would like to continue with university.

[22] Attrition in the administrative data increases over time and is due to a number of reasons. For very few students, lower-secondary teachers did not know whether students had requested the official transcripts to enroll in high school. For others, we could not determine which school they had enrolled in (if any). And finally, during high school, some students transferred to other schools, and we were not able to verify where. To minimize attrition in the administrative data, we cross-checked all missing students with official records of all students who participated in the Grade 12 graduation exam available at province level. For a student not to appear on those lists, they would have had to drop out of school, to transfer to a school in a different province, or to repeat the school year. As repetition is quite low, and we observe that only very few students switch schools in the first two years of high school, the most likely outcome is that these students dropped out. Nevertheless, we treat these observations as missing in our main specifications.

**Table 1**
Balance table: Pre-intervention characteristics in treatment and control schools.

| Variable | All | | | Treatment | | | Control | | | Treat. - Contr. | | Norm. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Mean | SD | N | Mean | SD | N | Mean | SD | Diff. | p-val | Diff. |
| Student Characteristics - Admin. Data | | | | | | | | | | | | |
| Female | 1715 | 0.54 | 0.50 | 862 | 0.54 | 0.50 | 853 | 0.53 | 0.50 | 0.01 | 0.60 | 0.02 |
| Age | 1715 | 15.05 | 1.32 | 862 | 15.11 | 1.36 | 853 | 15.00 | 1.28 | 0.00 | 0.97 | 0.08 |
| Distance to school (km) | 1715 | 3.55 | 3.78 | 862 | 4.01 | 3.95 | 853 | 3.09 | 3.54 | 1.15 | 0.03 | 0.25 |
| Distance to district town (km) | 1715 | 11.33 | 7.44 | 862 | 9.80 | 6.40 | 853 | 12.88 | 8.08 | −2.84 | 0.13 | −0.42 |
| Distance to high school (km) | 1715 | 9.68 | 6.80 | 862 | 9.24 | 6.43 | 853 | 10.13 | 7.12 | −1.10 | 0.49 | −0.13 |
| Pre-grade, main subjects (standardized) | 1715 | 0.01 | 0.98 | 862 | −0.22 | 0.96 | 853 | 0.24 | 0.95 | −0.44 | 0.00 | −0.48 |
| Avg absence (Dec&Jan) | 1715 | 1.52 | 2.02 | 862 | 1.63 | 1.92 | 853 | 1.41 | 2.10 | 0.30 | 0.20 | 0.11 |
| School Characteristics - Admin. Data | | | | | | | | | | | | |
| Class Size | 1715 | 45.71 | 11.04 | 862 | 46.15 | 11.52 | 853 | 45.26 | 10.52 | 1.90 | 0.49 | 0.08 |
| Teacher: Female | 1715 | 0.33 | 0.47 | 862 | 0.29 | 0.45 | 853 | 0.36 | 0.48 | −0.05 | 0.68 | −0.15 |
| Teacher: Age | 1715 | 32.42 | 6.53 | 862 | 29.86 | 5.34 | 853 | 35.00 | 6.61 | −4.37 | 0.02 | −0.86 |
| Teacher: Has University Degree | 1715 | 0.51 | 0.50 | 862 | 0.55 | 0.50 | 853 | 0.47 | 0.50 | 0.04 | 0.77 | 0.16 |
| Teacher: Log Distance to School (km) | 1715 | 1.57 | 1.22 | 862 | 1.80 | 1.17 | 853 | 1.34 | 1.23 | 0.47 | 0.12 | 0.38 |
| High school attached | 1715 | 0.17 | 0.37 | 862 | 0.14 | 0.35 | 853 | 0.19 | 0.39 | −0.02 | 0.89 | −0.14 |
| Partnership with Child's Dream | 1715 | 0.77 | 0.42 | 862 | 0.86 | 0.35 | 853 | 0.68 | 0.47 | 0.14 | 0.18 | 0.44 |
| Parental Characteristics - Phone Survey | | | | | | | | | | | | |
| Father completed ≤ primary educ. | 1170 | 0.81 | 0.39 | 581 | 0.84 | 0.37 | 589 | 0.79 | 0.41 | 0.04 | 0.31 | 0.13 |
| Mother completed ≤ primary educ. | 1246 | 0.92 | 0.27 | 622 | 0.93 | 0.26 | 624 | 0.92 | 0.27 | 0.01 | 0.53 | 0.04 |
| Any parents is farmer | 1327 | 0.69 | 0.46 | 666 | 0.70 | 0.46 | 661 | 0.68 | 0.47 | 0.00 | 0.99 | 0.04 |

*Notes:* Treatment - Control difference, and p-values are obtained by regressing variable of interest on a treatment dummy and district fixed effects with standard errors clustered at the school level. The pre-grade in the main subjects is the sum of Math, English and Khmer, averaged over the months December and January, and standardized across schools. The highest achievable points in Math, Khmer, and English are 100, 100 and 50, respectively. Absences are absent days per month. One school did not report absences, for this school the sample mean is imputed.

control students (as shown in Figure A.1), we primarily address these imbalances through standard regression-based corrections and data-driven approaches (double machine learning). As additional robustness check, we also exclude the schools in the tails of the grades distribution (class-averaged) from the sample.[23]

We track student outcomes after the intervention through phone-survey answers, as well as administrative data obtained from lower-secondary and from high schools (as summarized in Table A.3). In terms of students' activities during school closure, 43% of the students strongly agreed with the statement "I kept studying during school closure", while only 25% reported that their main activity in the last 7 days was studying. Students' aspirations are quite high: 13.5 years of schooling on average. The vast majority (96%) reported to aspire to complete at least higher-secondary education and 44% aspired to a university degree. Similarly, a very large proportion aimed for a career that requires higher secondary education (91%) and about one out of three students aspired a career that requires a university degree. As a prize for participating in the phone survey, the great majority of students (80%) chose educational mentoring over phone credit.

Of the surveyed students, 24% reported to have applied for some kind of high-school related scholarship. This could be either the scheme operated by Child's Dream, which was available at 76% of the schools in our sample, a scholarship provided by the Cambodian government, or other scholarships from NGOs operating locally. From the administrative records of CD, we can infer that of all the students who had the possibility to apply for a CD scholarship (n=1317), 17% applied for it, and 4% received it.

We use students' participation in the final exam of Grade 9 as the main indicator for completion of the academic year. Of our targeted students, 13% did not attend the exam and are therefore considered dropouts. Among those who participated in the exam, most students performed surprisingly well. Students usually need 260 points to pass the exam and to be allowed to enroll in high school, and the vast majority exceeded this threshold.

Of all students, 81% requested their official transcripts for Grade 9, which are necessary in order to be able to enroll in high school. A smaller share of students (75%) actually started high school as confirmed by high school teachers. An even smaller share of students (64%) started Grade 11 a year later, and only 61% of all students who we were able to track attended the midterm exam of Grade 11, which 49% passed in the first attempt. Finally, about 54% of students were able to complete high school about three years and nine months after our intervention took place.

## 4. Empirical analysis

### *4.1. Empirical approach*

In order to analyze whether our intervention affected schooling decisions, we focus on six outcomes that are self-reported and based on answers collected in the phone survey and on six outcomes constructed from administrative data. With the phone-survey data, we analyze whether students strongly agreed with the statement that they kept studying during the first school closure, as well as students' aspirations, in particular, the years of schooling students aspire to achieve, whether the job they would like to do when they are 25 is outside the typical reference window (*i.e.*, it is not teacher, general practitioner, police officer, or soldier), and the years of schooling that the aspired occupation requires. In addition, we consider whether students rate the probability that they will achieve their stated educational (and occupational) aspiration as high.[24] Using administrative data, we study whether students completed Grade 9 (*i.e.*, they participated in the final exam), their performance in the final exam (standardized across schools), as well as whether they enroll in high school. To study longer-term effects, we investigate whether students started Grade 11, whether they attended the Grade 11 midterm exam, and whether they

[23] It is important to note that the baseline imbalance can entirely be explained by the COVID-19 related interruptions to the implementation. Had we been able to implement the intervention in all 30 treatment schools, the normalized difference in pre-intervention grades would have been 0.12, and the *p*-value on the Treatment - Control difference would have been 0.58.

[24] These two variables are coded as dummies that take the value one if the stated confidence on a scale from zero to 10 is greater than 5. We use the dummy rather than the continuous variable because of substantial bunching in the data around the middle value.

completed high school. Of each group of outcomes, the first three were pre-specified in the pre-analysis plan (Gehrke et al., 2020), while the latter three serve to provide deeper insights into the mechanisms at play or to understand the persistence of the estimated effects.

Because not all students were present during the intervention, we estimate both intention-to-treat effects (ITT) as well as treatment-on-the-treated effects (TOT). ITT estimates rely on the original treatment assignment, *i.e.*, whether a student is enrolled in a school and class in which the workshop was conducted. The specification takes the form:

$$Y_{ijd} = \alpha + \beta T_j + \gamma' X_{ijd} + \xi_d + \epsilon_{ijd}, \tag{1}$$

where $Y_{ijd}$ is any of the outcomes of interest for student $i$ in school $j$ and district $d$. $T_j$ is a dummy equal to one if the intervention was implemented in school $j$ and zero otherwise. $X_{ijd}$ is a vector of pre-specified student and school characteristics (student age, gender, pre-intervention grades and absence, class size, and CD partnership), and $\xi_d$ are district fixed effects. $\epsilon_{ijd}$ is the idiosyncratic error term. Standard errors are corrected for clustering within schools.

The TOT estimates take into account that not all targeted students actually attended the workshop (in total 79 of the 862 students that were targeted did not show up on the day of the workshop). To address this, we estimate Eqs. (2) in 2SLS, instrumenting $Treated_i$ with the original treatment assignment ($T_j$).

$$Treated_i = \eta + \theta T_j + \kappa' X_{ijd} + \phi_d + \nu_{ijd} \tag{2a}$$

$$Y_{ijd} = \alpha + \beta \widehat{Treated}_i + \gamma' X_{ijd} + \xi_d + \epsilon_{ijd}. \tag{2b}$$

In addition, we use a data-driven approach, namely double machine learning (DML, Chernozhukov et al., 2018), to select the relevant covariates and estimate the coefficients of interest.[25] The double machine learning method follows the partialing-out lasso, yet tunes the parameters of the lasso via cross-validation. It is generally considered the most suitable solution for lasso-based inference (Baiardi and Naghi, 2024; Cameron and Trivedi, 2022).

### 4.2. Main results

Results are presented in Tables 2 and 3. We report ITT effects (Panel A) and TOT effects (Panel B). For each outcome, we first show the estimates that control for student and school characteristics as well as district fixed effects as pre-specified. We then present the double machine learning estimates. We depict standard errors clustered at the school level in parentheses, and Anderson's (2008) sharpened q-values to account for the False Discovery Rate (FDR) in brackets. Our preferred specification is shown in the even columns of panel B, and we base inference on the sharpened q-values.

Using phone-survey data, which were collected about five months after the intervention, we find that the intervention had no effect on average on whether students self-report to have been studying during the lockdown period (Table 2). We do, however, find a weakly positive effect on students' educational aspirations in years of schooling (0.24 additional years), as well as a weakly negative effect on the diversification in students' occupational aspirations (−5 pp), which implies that treated students were less likely to name an occupation other than the main four as their stated career goal. This result is strongly driven by the higher likelihood that treated students mentioned 'teacher' as their occupational aspiration, which had been featured as part of the career exploration tool. When studying occupational aspirations in terms of the years of schooling required to carry out the aspired occupation, we find a positive effect that is comparable in magnitude to the effect observed on educational aspirations. Overall, treated students were slightly more confident in being able to attain the aspired level of education or the aspired occupation. However, none of the effects discussed above are statistically significant after correcting for the FDR.[26]

In terms of administrative outcomes (Table 3), we find that treated students were 2.6 percentage points more likely to return to school and attend the Grade 9 final exam after the COVID-19 lockdown (3% increase over the control group mean). Conditional on attending, treated students also performed better than control students in the final exam (0.22 SD). These two effects are statistically significant after correcting for the FDR only in the DML specification. The magnitude of the effect on final exam performance is striking given that the intervention also led to higher attendance in that exam, and the marginal students to attend the exam were likely among the students with lower school performance.[27] In addition, we find a positive effect of 5.9pp (7.9%) on high school enrollment. This effect is statistically significant at the 10% level after correcting for the FDR in both specifications, and slightly larger (9.6pp) when using OLS with pre-specified controls for estimation.

This positive effect on enrollment is sustained throughout high school. Treated students were 6.2pp (9.7%) more likely to start Grade 11 and 5.6pp (9.2%) more likely to attend the midterm exam of Grade 11. Again the OLS estimates are slightly larger than the DML estimates, but less precisely estimated. All coefficients remain statistically significant after correcting for the FDR. Finally, treated students were 3.9pp (7%) more likely to complete high school. Although the effect size is somewhat attenuated (potentially because of the larger number of missing observations), the effect is statistically significant at conventional levels after correcting for the FDR.

Taken together, these results imply that participation in the workshop encouraged students to increase their educational investments. The estimated effect sizes on enrollment and graduation are comparable in magnitude to those found in other studies of information or career guidance interventions. Dinkelman and Martínez A. (2014) show that providing Chilean eighth graders with a short informational DVD about financial aid increases college preparatory high school enrollment by 6pp. Hoest et al. (2013) find that a reform of the career guidance system for lower secondary students in Denmark increased upper secondary school enrollment by 4.0–6.3pp among immigrant students. Carlana et al. (2022) show that tutoring and career counseling provided to high-ability immigrant students in Italy increases enrollment in an academically-oriented high school track by approximately 4pp, while Renée (2025) finds a 10pp increase in college enrollment from a career guidance program in Canada. Students in rural Cambodia had almost certainly never received any form of career guidance prior to our intervention, let alone guidance matched to their individual interests, which likely explains why even a half-day workshop can have a substantial impact at the margin.

### 4.3. Heterogeneities by academic performance

Given that the workshop was designed to help students formulate aspirations that are more aligned with their abilities and interests, the intervention's effect on aspirations and educational investment might

[25] We use a lasso-based procedure, sometimes referred to as cross-fit partialing-out lasso. We implement the procedure in Stata using `xporegress` and `xpoivregress`. For the cross-fitting, we report results averaged over ten sample splits, for each using ten folds.

[26] The number of observations varies by survey outcomes as students could refuse to answer individual survey items. We test the sensitivity of our results to these missing values using bound analyses (see Section 4.4).

[27] This raises suspicion that teachers may have become more lenient due to our intervention. However, we do not find a statistically significant effect on the probability of passing the final exam (see Table A.4). This attenuates concerns that the observed effect on grades may be driven by teachers rather than students.

**Table 2**
Main results – phone-survey data.

| | Studied during lock down | | Educ aspirations (years schooling) | | Occup aspirations (outside ref window) | | Occup aspirations (years schooling) | | Believes to reach aspired educ | | Believes to reach aspired occup | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *Panel A: Intention to Treat* | | | | | | | | | | | | |
| Treatment Assigned | 0.049 | 0.008 | 0.127 | 0.224 | −0.064 | −0.046 | 0.252 | 0.217 | 0.050 | 0.012 | 0.064 | 0.033 |
| | (0.041) | (0.028) | (0.133) | (0.121)* | (0.029)** | (0.024)* | (0.135)* | (0.107)** | (0.035) | (0.030) | (0.031)** | (0.028) |
| | [ 0.169] | [ 0.582] | [ 0.209] | [ 0.153] | [ 0.161] | [ 0.153] | [ 0.161] | [ 0.153] | [ 0.161] | [ 0.582] | [ 0.161] | [ 0.226] |
| *Panel B: Treatment on the Treated* | | | | | | | | | | | | |
| Treated | 0.053 | 0.009 | 0.138 | 0.242 | −0.068 | −0.049 | 0.271 | 0.233 | 0.054 | 0.013 | 0.069 | 0.036 |
| | (0.043) | (0.031) | (0.141) | (0.132)* | (0.031)** | (0.026)* | (0.142)* | (0.115)** | (0.037) | (0.033) | (0.033)** | (0.031) |
| | [ 0.155] | [ 0.578] | [ 0.197] | [ 0.153] | [ 0.127] | [ 0.153] | [ 0.127] | [ 0.153] | [ 0.127] | [ 0.578] | [ 0.127] | [ 0.224] |
| OLS w pre-specified controls | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | |
| Cross-fit partialing out Lasso | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Control Mean | 0.42 | | 13.45 | | 0.23 | | 12.80 | | 0.43 | | 0.37 | |
| Observations | 1296 | | 1317 | | 1291 | | 1291 | | 1291 | | 1289 | |

*Notes:* Panel A: OLS estimates. Treatment Assigned = The student's school received the intervention. Panel B: 2SLS estimates. Treatment (participation in the intervention) instrumented with assigned treatment status. Standard errors (in parentheses) are clustered at the school level. In brackets, Anderson's sharpened q-values to account for the False Discovery Rate (Anderson, 2008). In the uneven columns, we report results from OLS with pre-specified controls including student age, gender, pre-intervention grades and absence, class size, Child's Dream partnership, and district fixed effects. In the even columns, we report results from cross-fit partialing out lasso using 10 folds and 10 re-samples. */**/*** denote significance levels at 10/5/1 percent respectively.

**Table 3**
Main results - administrative data.

| | Attended grd 9 final exam | | Final exam grade (std.) | | High school enrollment | | Started grd 11 | | Attended grd 11 midterm exam | | Graduated high school | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| *Panel A: Intention to Treat* | | | | | | | | | | | | |
| Treatment Assigned | 0.040 | 0.024 | 0.156 | 0.200 | 0.088 | 0.054 | 0.086 | 0.057 | 0.088 | 0.051 | 0.064 | 0.036 |
| | (0.028) | (0.018) | (0.151) | (0.052)*** | (0.036)** | (0.023)** | (0.043)* | (0.025)** | (0.041)** | (0.025)** | (0.041) | (0.026) |
| | [ 0.117] | [ 0.066] | [ 0.183] | [ 0.001] | [ 0.117] | [ 0.042] | [ 0.117] | [ 0.042] | [ 0.117] | [ 0.051] | [ 0.117] | [ 0.066] |
| *Panel B: Treatment on the Treated* | | | | | | | | | | | | |
| Treated | 0.044 | 0.026 | 0.167 | 0.215 | 0.096 | 0.059 | 0.094 | 0.062 | 0.096 | 0.056 | 0.070 | 0.039 |
| | (0.031) | (0.020) | (0.159) | (0.056)*** | (0.039)** | (0.025)** | (0.046)** | (0.028)** | (0.044)** | (0.027)** | (0.044) | (0.028) |
| | [ 0.102] | [ 0.066] | [ 0.171] | [ 0.001] | [ 0.085] | [ 0.042] | [ 0.085] | [ 0.042] | [ 0.085] | [ 0.051] | [ 0.093] | [ 0.066] |
| OLS w pre-specified controls | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | |
| Cross-fit partialing out Lasso | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Control Mean | 0.87 | | 0.04 | | 0.75 | | 0.64 | | 0.61 | | 0.56 | |
| Observations | 1,715 | | 1,485 | | 1,697 | | 1,693 | | 1,692 | | 1,672 | |

*Notes:* Panel A: OLS estimates. Treatment Assigned = The student's school received the intervention. Panel B: 2SLS estimates. Treatment (participation in the intervention) instrumented with assigned treatment status. Standard errors (in parentheses) are clustered at the school level. In brackets, Anderson's sharpened q-values to account for the False Discovery Rate (Anderson, 2008). In the uneven columns, we report results from OLS with pre-specified controls including student age, gender, pre-intervention grades and absence, class size, Child's Dream partnership, and district fixed effects. In the even columns, we report results from cross-fit partialing out lasso using 10 folds and 10 re-samples. */**/*** denote significance levels at 10/5/1 percent respectively.

differ by students' ability. We investigate this idea further by analyzing treatment effect heterogeneities along students' baseline academic performance.[28]

Figs. 3 and 4 report semi-parametric estimates (local mean smoothing) of each outcome variable at all levels of students' performance prior to the intervention (standardized sum of grades), separately for students in treatment and control schools. Each of these estimates controls linearly for student and school characteristics as well as district fixed effects (corresponding to the odd columns in Tables 2 and 3). Parametric estimates that interact the treatment indicator with students' baseline performance and account both for pre-specified controls as well as lasso-selected controls are reported in Tables A.5 and A.6, and reveal the same pattern.

We find strong evidence for heterogeneous treatment effects. Low-performing treated students were less likely to study during school closure than similarly performing students from control schools. By contrast, students who had been performing better than the median student before the intervention seem to have benefited from it; these students were significantly more likely to be studying at the time of the phone survey. In terms of educational aspirations, we find evidence that low-performing students downward adjusted their educational aspirations, while high-performing students adjusted them upwards. A similar pattern —albeit somewhat less pronounced among low-performing students— can be observed for occupational aspirations when expressed in years of schooling.[29] By contrast, we find no effect on the diversification in career goals for low-performing students, yet a negative effect for high-performers. This effect is driven by high-performing treated students being more likely to state 'teacher' as their occupational aspiration. As the intervention features high-school teacher as occupation, students may have positively updated on the attractiveness of this occupation. We find no clear pattern by baseline academic performance with respect to student's beliefs about whether they can reach these aspirations. In terms of beliefs that students will

[28] We have not pre-specified any heterogeneity analysis, and this section is exploratory by nature. We find no treatment effect heterogeneities by gender or parental background (results available on request).

[29] The interaction effect is statistically significant for this outcome in the DML specification but not in OLS (see Table A.5). This attenuation may be due to the fact that we manually coded years of schooling for all listed occupations, which may not always be correct.

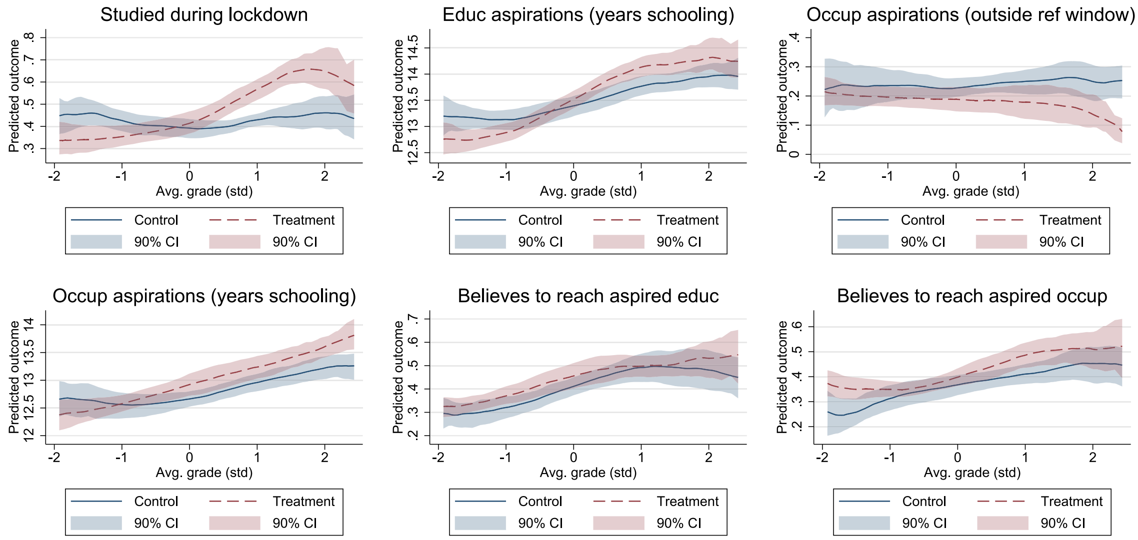


**Fig. 3.** Treatment effect heterogeneity by pre-intervention Grades (1).
*Notes*: These figures display the weighted moving-average (bandwidth = 0.5, Epanechnikov kernel) and bootstrap confidence intervals (clustered at the level of the school) of the dependent variable (in the figure header) over the baseline grade (sum of Math, Khmer and English grades, averaged over the months December and January, and standardized over all schools), separately for treatment and control schools. To produce semi-parametric estimates, the dependent variable is first partialled out from District fixed effects, as well as student controls (gender, age, average absence) and school controls (class size, Child's Dream partnership).

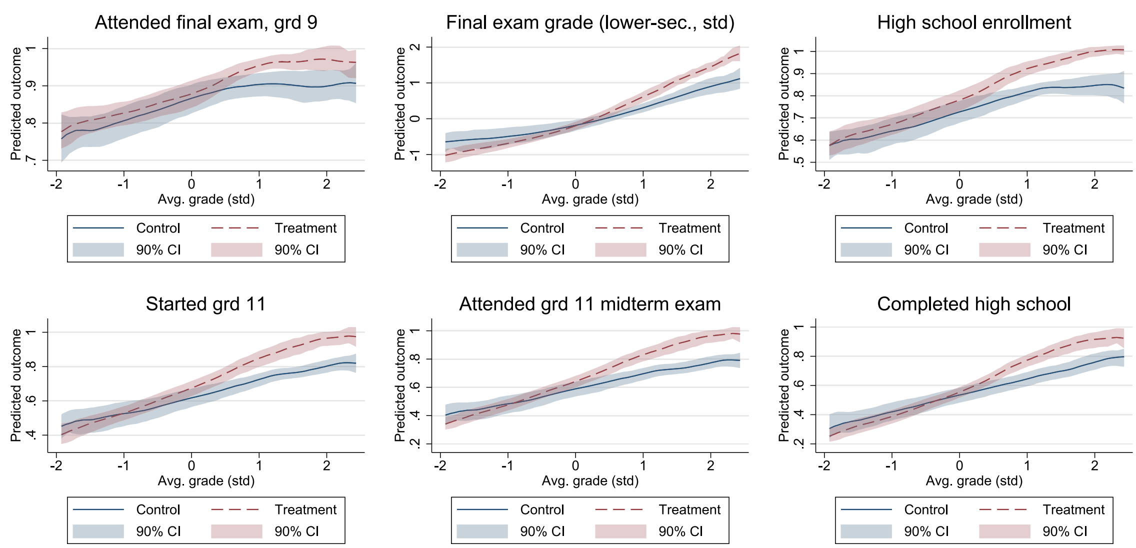


**Fig. 4.** Treatment effect heterogeneity by pre-intervention Grades (2).
*Notes*: These figures display the weighted moving-average (bandwidth = 0.5, Epanechnikov kernel) and bootstrap confidence intervals (clustered at the level of the school) of the dependent variable (in the figure header) over the baseline grade (sum of Math, Khmer and English grades, averaged over the months December and January, and standardized over all schools), separately for treatment and control schools. To produce semi-parametric estimates, the dependent variable is first partialled out from District fixed effects, as well as student controls (gender, age, average absence) and school controls (class size, Child's Dream partnership).

reach the aspired level of education, the intervention seems to have had marginal but positive effects over the entire distribution of initial grades. With respect to beliefs about reaching the aspired occupation, the effects are again positive throughout, and most pronounced among the lowest-performing students and those that perform somewhat better than average.

Treatment effects are also heterogeneous with respect to participation in the final exam, and more strongly so for students' performance in that exam; low-performing treated students were marginally more likely to participate in the final exam, but did perform worse than their control-school counterparts conditional on participating. High-performing treated students, in turn, were (up to 10pp) more likely to participate in the final exam and also performed better than their counterparts from control schools.[30]

The positive average effect on high school enrollment seems to be entirely driven by better-performing students, for whom we find substantially higher high-school enrollment rates than for similarly

[30] Note that this positive effect on high performers is unlikely to be driven by selection on performance, as we are flexibly conditioning on baseline academic performance. It may still be the case, however, that our intervention motivated a very specific group of students within the high performers to participate in the exam (e.g., students with higher socio-emotional skills), which may drive the positive effect on final exam grades.

high-performing students from control schools. Among low-performing students, by contrast, high-school enrollment is not differentially affected. This can be explained by the fact that high-school enrollment rates are relatively low among this group. While more than 75% of the low-performing students participated in the final exam, only about 60% enrolled in high school. In other words, any negative effect on low-performing students disappears about nine months after the intervention, as a large fraction of students performing in the lower-end of the grade distribution would have dropped out by that time anyway. The positive average effect throughout high school continues to be driven by high-performing students, who were significantly more likely to start Grade 11, to attend the midterm exam of Grade 11, as well as to complete high school than high-performing students from control schools.

In summary, we find considerable heterogeneities in treatment effects by pre-intervention academic performance. The negative effect on low-performing students seems to vanish at the end of lower-secondary school, as low-performing students were not less likely to enroll in high school. This suggests that the intervention led low-performing students to decide sooner against attending high school, and that low-performing students adjusted their effort accordingly. Consistent with that interpretation, we find that low-performing treated students (especially those with a standardized grade below minus one) were less likely to report in the phone survey that their main activity in the last 7 days was studying, and to select educational mentoring (rather than phone credit) as a prize for participating in the phone-survey (for these and the following results see Figure A.2 and Table A.4). Low-performing students were also less likely to apply for a scholarship as self-reported in phone survey, and as confirmed by administrative data from CD. However, they were *not* less likely to pass the final exam of Grade 9 nor to obtain a CD scholarship, suggesting that low-performing treated students invested just enough effort in education to be able to graduate lower-secondary school, while their control school counterparts studied more during Grade 9, but still did not start high school.

### 4.4. Robustness checks

We perform a number of robustness checks to corroborate the main findings, as well as the heterogeneities by student performance.

*First*, we investigate whether inference is sensitive to alternative assignments of treatment status in our sample or to how we correct for multiple hypothesis testing. The p-values on the main effects (on high school enrollment, Grade 11 enrollment and performance, and high school completion) as well as the treatment effect heterogeneities (on study behavior, on educational and occupational aspirations, and on schooling outcomes) remain similar irrespective of whether we use randomization inference or control for the family-wise error rates using the Romano-Wolf stepdown procedure (see Table A.7).

*Second*, we evaluate the robustness of our findings to selective attrition. For the survey-based outcomes, we re-weight the sample with the inverse of the probability of participating in the phone survey. We predict phone-survey participation with student, teacher, and school characteristics obtained from administrative data as described in Online Appendix C. Our main findings, as well as the treatment effect heterogeneities by pre-intervention student performance, are unchanged (see Table A.8). Following the strategy outlined in Kling et al. (2007), we also calculate sensitivity bounds for our estimates by varying the assumptions about outcomes for those students that did not answer individual survey questions in the phone survey or that could not be tracked throughout high school. We start with very extreme assumptions, *i.e.*, setting the outcome values of attritors to the minimum value for students in the treatment arm and to the maximum value for students in the control arm, and then relax these assumptions step-by-step. For phone-survey outcomes, the interaction effects on studying during the lockdown and educational aspirations are significant throughout, while the interaction effect on occupational aspirations (in years of schooling) declines somewhat in magnitude, which can likely be attributed to the larger number of missing values in this variable. For high school-related outcomes (enrollment in high school and in Grade 11, attendance of the Grade 11 midterm exam, high school completion), the average treatment effects, as well as the treatment effect heterogeneities are largely robust. Only the average treatment effect on high school completion is insignificant in the specifications in which more extreme assumptions are taken (see Tables A.9 and A.10). However, it is worth keeping in mind that most of the students we could not track probably have dropped out of high school (column (8) shows that our estimates are robust to this assumption too).

*Third*, we address remaining concerns regarding the imbalance in baseline grades, and the possibility that it could be driving the observed effects, given that baseline grades are strong determinants of subsequent schooling outcomes. We first identify which schools drive these imbalances. We find that treatment schools are substantially over-represented among the schools with the lowest average grades and under-represented among the schools with the highest grades.[31] If we restrict the sample to the schools above the 10th and below the 90th percentiles in average grades, our sample drops to 1298 students from 13 treatment and 14 control schools. As shown in Table A.11 baseline characteristics are mostly balanced in this subsample, and the normalized difference in baseline grades is −0.24. We then estimate the main treatment effects as well as treatment effect heterogeneities by baseline grades in this subsample, and find that our results are virtually unchanged (see Tables A.12 and A.13).[32]

*Fourth*, we study whether the observed treatment effect heterogeneities by performance are mere spurious correlations driven by confounding variables that affect pre-intervention performance and educational investments. To identify potential confounders, we employ two strategies. We first analyze potential correlates of academic performance by regressing an extensive set of household-level and school-level variables on academic performance while controlling for student's age, gender, and district fixed effects. For this we use administrative data and information from the phone survey. At the household level, neither remoteness of the student's village (in terms of distance to the school or next district town) nor occupation of the parents in agriculture are significantly correlated with academic performance. However, we find that students' pre-intervention grades are correlated with parental education, and with the predicted probability (based on the sector of occupation) that one of the parents lost their job or income during the COVID-19 crisis (c.f. Table A.14).[33] At the school level, we find that teacher age is correlated with baseline performance, but none of the other characteristics: class size, gender of teacher, teacher distance to school, whether a high school is attached to the lower-secondary school and whether the school partners with CD (c.f. Table A.15). We then split our sample into four quartiles of student baseline academic performance and investigate which student, school, or parental characteristics are ever imbalanced in each of these quartiles.[34] For each of the potential confounders (that are either significantly correlated with academic performance or that ever show up as imbalanced), we then add the interaction of the confounder with treatment

[31] We also investigate if these imbalances are driven by particular districts, and find that they materialize in almost all of the sample districts.

[32] In each of these regressions, we control for district fixed effects, as well as student age, gender, and baseline grades, in addition to the teacher age, the distance the teacher lives from the school, whether the school has a high school attached, and whether the school partners with CD.

[33] In previous work (Gehrke et al., 2023), we show that the economic repercussions of the COVID-19 crisis affected students' academic performance mostly through parental income losses.

[34] The variables that show up as imbalanced at the 10% level at least once are: student gender, distance to school, high school, and to district town, teacher age, and distance to school (log), and school partnership with CD.

status to the main regression. For parental characteristics, we can only carry out this exercise for the students who participated in the phone survey and reported information on their parents' education or job loss; we consider these a selected sample. The analysis is therefore split in two. We first interact treatment with parental characteristics (in Tables A.16 and A.17).[35] Overall, our results are unchanged. In a separate exercise, we add interactions of any other individual- or school-level characteristic identified as potential confounders with treatment (in Tables A.18 and A.19). Again our results are unchanged, suggesting that none of these student, parental, teacher, or school characteristics are driving the observed heterogeneities in treatment effects.

*Fifth*, we investigate if the heterogeneities we find are merely driven by differences between treatment and control schools in how strongly baseline grades predict subsequent outcomes.[36] This seems not to be the case, as the semi-parametric relationship between February grades (which we were able to collect from 17 treatment schools and 14 control schools) and the main running variable (grades averaged over December and January) are virtually identical (as shown in Figure A.3).[37]

### 4.5. Implications of COVID-19

The intervention took place at the onset of COVID-19, which affected the whole economy and brought about high levels of uncertainty. Although there were very few reported cases of COVID-19 infections in Cambodia in 2020, the global economic recession and travel disruptions had severe repercussions on households. In the phone survey conducted during the first school closure, many students reported that their parents lost income or even their job due to the crisis (see also Gehrke et al., 2023). Students were thus facing severe financial constraints, and over 60% of the students reported that they had to start working to support their family due to the crisis. The majority of students believed that the crisis lowered job prospects. An even larger share indicated being afraid of not being able to continue to Grade 10 due to the crisis, with most fearing their family would have fewer financial resources to support them going to school. The distribution looks very similar for students from treated and control schools as well as for low-performing and high performing students (see Figure A.4).

The crisis seems to also have reduced aspirations and confidence uniformly across all students. Using data on aspirations collected immediately after the intervention (for treatment schools only) and in the phone survey roughly five months later, we find that aspired years of schooling dropped substantially between these two surveys. This drop is uniform across the entire baseline performance distribution. A similar pattern holds for students' confidence in reaching their aspired education level (see Figure A.5). This period of high uncertainty in the first months after the first lockdown might help explain why we see limited impact of the intervention on aspirations and confidence, yet stronger effects on the administrative outcomes that were collected later. While our data do not allow us to speak to this directly, the workshop could have supported students in treatment schools in better coping with the disruptions.

## 5. Mechanisms

The evidence presented so far suggests that the workshop increased educational investments as well as educational and occupational aspirations for high-performing students, while it decreased educational investments (at least in the short-run) as well as educational and occupational aspirations for low-performing students. To rationalize this treatment effect heterogeneity *post hoc*, we develop a model of endogenous aspiration formation under information constraints, which is able to capture the main patterns observed in the data. Nevertheless, we cannot fully rule out the importance of alternative explanations, which we discuss further below.

### 5.1. Endogenous aspirations and information constraints

Consider a model of endogenously determined aspirations as in Dalton et al. (2016), and assume that students choose educational aspirations and occupational aspirations that are aligned with each other. For example, a student may want to become a medical practitioner and knows that they need to get a university degree (in medicine) to achieve this. Imposing that occupational and educational aspirations are aligned with each other implies that students' aspirations can be formulated over educational outcomes, even if students do not derive utility from achieving any given level of education *per se* but rather from the occupational opportunities that arise from this achievement. The utility function of a student can then be described by:

$$u = b(\theta) + v\left(\frac{\theta - a}{\theta}\right) - c(e), \tag{3}$$

with $\theta$ being the final level of education, $a \in R_+$ being the aspired level of education, and $e \in [0,1]$ being effort. We follow Dalton et al. (2016) in our assumptions about the functional forms of the three additive components of $u$. We assume that direct utility, $b(\theta)$, is concave and twice differentiable with $b(0) = 0$, and the coefficient of relative risk aversion $r(x) = -\theta b''(\theta)/b'(\theta) < 1$. Milestone utility, $v[(\theta - a)/\theta]$, then, is continuously differentiable with $v'(0) > 0$ and $[v'(x) - v''(x)(1-x)] \geq 0$ for all feasible values of $x$, with one possible formulation being an S-shaped value function with diminishing sensitivity as in Kahneman and Tversky (1979). The functional form of $v(x)$ ensures that the utility associated with satisfying any aspiration increases in the level of $a$. The cost of effort, $c(e)$, finally, is continuous, twice differentiable and convex in $e$, with $c(0) = 0$.

When choosing educational aspirations, students take into account their own preferences, while being constrained by their academic ability, *i.e.*, the rate at which educational investments transform into educational outcomes. Similarly to Dalton et al. (2016), we incorporate these external constraints by imposing that skills $\mu \in [0, \infty]$ and effort $e$ are complements in the production of education $\theta$. For simplicity, assume:

$$\theta = f(\mu, e) = (1 + \mu)e. \tag{4}$$

A rational student chooses the optimal effort-aspiration pair $(e^*, a^*)$ that maximizes utility. As in Dalton et al. (2016), we constrain possible solutions to consistent effort-aspiration pairs, *i.e.*, pairs in which the aspiration equals the final outcome: $a = f(\mu, e) = (1 + \mu)e$.[38] This assumption, together with Eq. (4) and the functional form assumptions concerning $u$ are sufficient to ensure a unique solution level of aspiration and effort, which is increasing in skills. The association between aspirations and skills in optimum is weakly positive throughout and strictly positive for interior solutions. Intuitively, each student sets their aspirations to be high enough to incentivize effort, but not too high (as they otherwise become unattainable and lead to frustration). However, in our context, we find that educational and occupational

[35] In the even columns, we report our main OLS specification with interaction effects for the selected sample, in the odd columns we include parental education, parental job loss, and parental income loss interacted with treatment status.

[36] A reason may be, for example, the extent to which teachers concentrate their efforts towards high-performing students, or conversely, those who are falling behind.

[37] Note that it is unlikely for February grades to be affected by our treatment. While we visited nine schools in February, these visits were on the last days of the month, and monthly grades typically reflect performance throughout the month. Even if students responded by immediately increasing or reducing their effort, it is not clear that this would be picked up by teachers in such a short time.

[38] See Eq. (2) in their paper.

aspirations are only modestly correlated with baseline academic performance in the control group. To rationalize why our intervention increases the correlation between educational aspirations and initial academic performance, we have to accommodate the possibility that information frictions constrain students in choosing the right level of aspirations. In line with Genicot and Ray (2017), we therefore assume that aspirations, although endogenously defined, are drawn from a distribution of known outcomes, the aspirations window. This implies that

$$a^* = \Psi(\mu, e^*, F), \tag{5}$$

where $F$ is the known distribution of education-occupation combinations. An attenuation in the relationship between innate abilities and educational aspirations can then arise for two reasons. *First*, students lack information about career possibilities and are constrained by the overly narrow set of occupations they know. *Second*, students misperceive the educational requirements associated with their occupational aspiration and mistakenly set their educational aspirations to the wrong level. Both possibilities can lead to a truncated aspirations window, and cause students to define educational aspirations that do not match their abilities —too high (for low-performers) or too low (for high performers)— resulting in a relatively flat relationship between baseline academic performance and aspirations.

Importantly, if aspirations are insufficiently responsive to innate ability, students may choose levels of investment that are individually sub-optimal. Some students may keep investing in education, even though their abilities do not match with higher education and they could develop occupational aspirations that do not require higher education, while others underinvest in education not knowing the true educational requirements for their occupational aspiration or not being aware of sufficiently ambitious aspirations that would match their ability.

Such a framework explains why providing students with the necessary tools to re-assess their occupational aspirations, while delivering information about the educational requirements necessary to carry out these jobs, helps students re-calibrate their aspirations, and adjust educational investments accordingly. In line with the predictions of our model, we observe a strong correlation between academic performance and types of occupations that students read about in the CET (see Table A.20). In particular, low-performing students spent more time reading about occupations that only require lower-secondary education.

### *5.2. Alternative explanations*

There are a number of alternative explanations for the observed effects. In the following, we discuss each of these mechanisms and the extent to which the empirical evidence supports them.

*Expectations.* Rather than broadening the aspirations window, the intervention may have forced students to adjust their expectations in heterogeneous ways. For example, the information on minimum requirements with respect to the grades students need to obtain in order to be able to enroll in high school, and the high standards applied to receive a scholarship, might have led low-performing students to realize that their academic performance does not match with a path of higher education. In contrast, for high-performing students, this same information may have helped them understand how well they are qualified for higher education and given them the necessary push to pursue this trajectory. However, we believe this mechanism is unlikely to explain the main findings for two main reasons. *First,* in data from the workshop midline survey, we already observe a strong gradient in self-reported beliefs about high school success by baseline grades (see Figure A.6). Information updating about high school success is therefore unlikely to be the driver of the observed effects on aspirations and schooling outcomes. *Second*, educational expectations do not adjust in exactly the same way as educational aspirations: while we do find that low-performing students downward adjusted their expectations, high-performing students (in particular those with standardized grade above 1) were not expecting more years of schooling than those in control schools (see bottom panel of Figure A.5), while their aspirations and investments clearly diverged. Together, these findings suggest that adjustments in expectations are unlikely to be the main driver underlying our results.

*Differential information acquisition.* An alternative explanation may be that different aspects of the information provided during the intervention were acquired and processed by low- vs. high-performers. For instance, low-performing students may only have absorbed and processed the information about high school costs, while high performers processed all information accurately, including costs, scholarship opportunities, etc. However, we find no evidence of differential cost updating for low- vs. high-performers among treated students. In fact, all students seem to update in the right direction (as shown in Figure A.7).[39]

*Discouragement.* Another alternative explanation may be that the intervention initially led all students to adjust their aspirations upwards, but that low-performing students quickly realized after the intervention that they are not able to attain the good grades required to enroll in high school, and became frustrated, thereby adjusting their aspirations to new, even lower, levels (as in McKenzie et al., 2022). However, we do find a strong gradient in educational aspirations with baseline grades already immediately in the aftermath of the intervention in treatment schools (see top panel of Figure A.5), which suggests that discouragement effects are unlikely to explain our findings.[40]

*Broadening of occupational interests.* A final possibility could be that our intervention simply broadened the set of occupations that students are potentially interested in, similar to the mechanism described in Belot et al. (2019). While this is consistent with increased educational investments among the high-performers, this argument is less well suited in explaining why low-performing students would react to the intervention by studying less during the lockdown, and perform worse during the final exam (conditional on attending).

## 6. Conclusions

This study provides experimental evidence that a half-day career guidance workshop designed to equip students with information about a range of occupations, and providing guidance on the next steps in the students' educational pathways, can improve students' educational outcomes. We show that the average positive effects mask substantial heterogeneity by students' pre-intervention school performance. Treated low-performing students were less likely to study during school closure compared to low-performing students in control schools, had somewhat lower educational and occupational aspirations, and performed worse in the final exam of Grade 9, yet they were not less likely to transition to high school. By contrast, treated high-performing students were more likely to study during school closure than their control group peers, had somewhat higher educational and occupational aspirations, performed better in the final exam, and were more likely to transition to, progress in, and complete high school.

Our findings suggest that the workshop, which is low-cost and easily scalable, improves the quality of educational decision-making by helping students more closely align their aspirations to their potential. Given the high dropout rates observed during high school in our

[39] In addition, cost estimates collected in the phone survey do not differ by treatment status nor by academic performance, neither for the total cost nor for the cost of extra classes, which most students mentioned as most expensive part of attending high school (see Table A.21).

[40] Note that aspirations reported right after the workshop were systematically higher than in the phone survey, which is likely due to the disruptive effect of the COVID-19 pandemic. We also cannot systematically compare treatment and control schools in this exercise, as this information was collected in treatment schools only.

sample, and in the country more generally, this intervention has the potential to generate substantial efficiency gains. This effect is akin to the productivity gains associated with improving workers' sorting along their comparative advantage as documented in previous work (Papageorgiou, 2014). While we cannot rule out that the intervention decreased human capital among some low-performing students as they oriented away from pursuing higher education, our findings suggest that they were *not* less likely to complete lower-secondary school than their control school peers, and by adjusting their educational aspirations, these students were potentially less frustrated throughout Grade 9. For high-performing students, on the other hand, our intervention provided the incentives and information necessary to exert more effort in lower-secondary school, allowing them to start high school at higher rates, and apparently better prepared. The fact that the positive effect of the intervention persisted more than three and a half years after the workshop had taken place suggests that this could indeed be a useful intervention to be implemented at scale.

It is important to emphasize that the average treatment effect of an intervention of this type will essentially depend on the underlying distribution of baseline academic performance in the student population. In settings in which many students are well prepared for higher education but still opt out, the average treatment effects would potentially be larger. On the other hand, in regions in which few students are equipped with the skills that are necessary to attend high school, a similar intervention may actually have no or even a negative effect on schooling outcomes.

Our study comes with a number of caveats. *First*, the study was (unintentionally) conducted during a very specific time period. About midway into our intervention, schools were closed for half a year due to the COVID-19 pandemic and students had to study on their own with little external support and with severe financial constraints as many parents had lost income. This might have undermined the positive effect of the intervention. *Second*, we were not able to track students once they left school. We therefore do not observe what students were doing after they dropped out, whether they started to work, what type of jobs they pursued, and whether they enrolled in vocational training, which would be important to understand the long run implications of the workshop on low-performing students. This is an avenue for future research.

## CRediT authorship contribution statement

**Esther Gehrke:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Friederike Lenel:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Claudia Schupp:** Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We thank the editor, Jing Cai, two anonymous referees, Rith Sarakk (PEPY Empowering Youth), Som San and Khlok Yem (Child's Dream), as well as Jenny Aker, Alexandra Avdeenko, Jana Cahlikova, Patricio Dalton, Veronika Geyer, Thomas Ginn, Till Grüneberg, Renate Hartwig, Elise Huillery, Clément Imbert, Anett John, Karlijn Morsink, Julianna Nielson, Ingrid Hoem Sjursen, Amma Panin, Eva Raiber, Lisa Spantig, Susan Steiner and Tavneet Suri for valuable feedback. The paper also benefited from comments by conference and seminar participants at Wageningen University, the University of Göttingen, ENS Lyon, Radboud University, UNU-MERIT, University of Groningen, RWI, GDE Conference 2021, 2023, GlaD Conference 2021, EEA congress 2021, SEEDEC 2021, and the C4ED seminar. We thank Aiko Schmeisser, Chhern Sreyneang, Chhly Chaktokrong, Doeun Sohin, Hong Sarith, Hong Sarou, Khann Rada, Khoeut Sochea, Kouth Sochampawatd, Kul Vandy, Phon Loem Bobon, Sin Chanita, Teung Seila and Touch Hean for excellent fieldwork assistance. This RCT obtained ethical approval by the Ethics Committee of the University of Göttingen (IRB approval date February 11, 2020, amendments approved on June 22, 2020) and was pre-registered with the AEA Trial Registry (Trial ID: AEARCTR-0005460). Funding by the German Research Foundation (RTG 1723) is gratefully acknowledged.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jdeveco.2026.103898.

## Data availability

The data and replication files are available online at: 10.5281/zenodo.21027280.

## References

Abbiati, G., Argentin, G., Barone, C., Schizzerotto, A., 2018. Information barriers and social stratification in higher education: evidence from a field experiment. Br. J. Sociol. 69 (4), 1248–1270.

Ahmed, H., Mahmud, M., Said, F., Tirmazee, Z., 2024. Encouraging female graduates to enter the labor force: Evidence from a role model intervention in Pakistan. Econ. Dev. Cult. Chang. 72 (2), 919–957.

Anderson, M.L., 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. J. Amer. Statist. Assoc. 103 (484), 1481–1495.

Asri, A., Asri, V., Hoeffler, A., 2025. Unlocking young women's potential? The impact of a low-cost career guidance program. J. Dev. Econ. 179, 103662.

Avitabile, C., de Hoyos, R., 2018. The heterogeneous effect of information on student performance: Evidence from a randomized control trial in Mexico. J. Dev. Econ. 135, 318–348.

Badini, S., Gehrke, E., Lenel, F., Schupp, C., 2026. Expanding horizons: A randomized controlled trial on adolescents' career information acquisition. Forthcom. in: J. Hum. Cap..

Baiardi, A., Naghi, A.A., 2024. The value added of machine learning to causal inference: evidence from revisited studies. Econom. J. 27 (2), 213–234.

Belot, M., Kircher, P., Muller, P., 2019. Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice. Rev. Econ. Stud. 86 (4), 1411–1447.

Benveniste, L., Marshall, J., Araujo, C.M., 2008. Teaching in Cambodia. The World Bank and Ministry of Education, Youth and Sport.

Berger, J., Pope, D., 2011. Can losing lead to winning? Manag. Sci. 57 (5), 817–827.

Bettinger, E.P., Long, B.T., Oreopoulos, P., Sanbonmatsu, L., 2012. The role of application assistance and information in college decisions: Results from the H&R Block FASFA experiment. J. Econ. Perspect. 127 (3), 1205–1242.

Bhan, P.C., 2020. Do role models increase student hope and effort? Evidence from India. University of Glasgow Working Paper.

Bjorvatn, K., Cappelen, A.W., Sekei, L.H., Sørensen, E.Ø., Tungodden, B., 2020. Teaching through television: Experimental evidence on entrepreneurship education in Tanzania. Manag. Sci. 66 (6), 2308–2325.

Cameron, A.C., Trivedi, P.K., 2022. Microeconometrics Using Stata, Volume 2. Stata Press.

Carlana, M., La Ferrara, E., Pinotti, P., 2022. Goals and gaps: Educational careers of immigrant children. Econometrica 90 (1), 1–29.

Chandler, D., 2007. A History of Cambodia, fourth ed. Westview Press, Boulder, Colorado.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., Robins, J., 2018. Double/debiased machine learning for treatment and structural parameters. Econom. J. 21 (1), C1–C68.

Dalton, P.S., Ghosal, S., Mani, A., 2016. Poverty and aspirations failure. Econ. J. (126), 165–188.

de Koning, B., Fouarge, D., Dur, R., 2025. Correcting Beliefs About Job Opportunities and Wages: A Field Experiment on Education Choices. IZA Discussion Paper No. 17951.

Dinkelman, T., Martínez A., C., 2014. Investing in schooling in Chile: The role of information about financial aid for higher education. Rev. Econ. Stat. 96 (2), 244–257.
Dizon-Ross, R., 2019. Parents' beliefs about their children's academic ability: Implications for educational investments. Am. Econ. Rev. 109 (8), 2728–2765.
Eng, S., Szmodis, W., Mulsow, M., 2014. Cambodian parental involvement. Elementary Sch. J. 114 (4), 573–594.
Gehrke, E., Lenel, F., Schupp, C., 2020. Trial registry for "Career goals and investments in education: Experimental evidence from Cambodia". AEARCTR-0005460.
Gehrke, E., Lenel, F., Schupp, C., 2023. COVID-19 Crisis, Economic Hardships, and Schooling Outcomes. Educ. Financ. Policy 18 (3), 522–546.
Genicot, G., Ray, D., 2017. Aspirations and inequality. Econometrica 85 (2), 489–519.
Genicot, G., Ray, D., 2020. Aspirations and economic behavior. Annu. Rev. Econ. 12 (1), 715–746.
Goux, D., Gurgand, M., Maurin, E., 2017. Adjusting your dreams? High school plans and dropout behaviour. Econ. J. 127 (602), 1025–1046.
Heckman, J.J., Mosso, S., 2014. The economics of human development and social mobility. Annu. Rev. Econ. 6 (1), 689–733.
Hoest, A., Jensen, V.M., Nielsen, L.P., 2013. Increasing the admission rate to upper secondary school: the case of lower secondary school student career guidance. Educ. Econ. 21 (3), 213–229.
Hoff, K.A., Song, Q.C., Wee, C.J., Phan, W.M.J., Rounds, J., 2020. Interest fit and job satisfaction: A systematic review and meta-analysis. J. Vocat. Behav. 123, 103503.
Holland, J.L., 1959. A theory of vocational choice. J. Couns. Psychol. 6 (1), 35–45.
Holland, J.L., 1997. Making Vocational Choices: A Theory of Vocational Personalities and Work Environments / John L. Holland, third ed. Psychological Assessment Resources, Odessa, Fla..
Hoxby, C.M., Turner, S., 2015. What high-achieving low-income students know about college. Am. Econ. Rev. 105 (5), 514–517.
Huang, H., Filmer, D., Fukao, T., 2017. Trends and Linkages in Schooling and Work among Cambodian Youth. World Bank.
ILO, 2024. Developing a Career Development Framework for Low- and Middle-Income Countries. International Labour Office, Geneva.
Jensen, R., 2010. The (perceived) returns to education and the demand for schooling. Q. J. Econ. 125 (2), 515–548.
Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47 (2), 263–291.
Kerr, S.P., Pekkarinen, T., Sarvimäki, M., Uusitalo, R., 2020. Post-secondary education and information on labor market prospects: A randomized field experiment. Labour Econ. 66, 101888.
Kipchumba, E.K., Porter, C., Serra, D., Sulaiman, M., et al., 2023. Influencing youths' aspirations and gender attitudes through role models: Evidence from somali schools.
Kling, J.R., Liebman, J.B., Katz, L.F., 2007. Experimental analysis of neighborhood effects. Econometrica 75 (1), 83–119.
La Ferrara, E., 2019. Presidential address: Aspirations, social norms, and development. J. Eur. Econ. Assoc. 17 (6), 1687–1722.
Leight, J., Gilligan, D.O., Mulford, M., Taffesse, A.S., Tambet, H., 2021. Aspiring to more? New Evidence on the Effect of a Light-Touch Aspirations Intervention in Rural Ethiopia. vol. 2070, Intl Food Policy Res Inst.
Lockwood, P., Kunda, Z., 1997. Superstars and me: Predicting the impact of role models on the self. J. Pers. Soc. Psychol. 73 (1), 91–103.
Mani, A., Riley, E., 2021. Social networks as levers of mobility. In: Iversen, V., Krishna, A., Sen, K. (Eds.), Social Mobility in Developing Countries: Concepts, Methods, and Determinants. Oxford University Press.
Mann, A., Denis, V., Schleicher, A., Ekhtiari, H., Forsyth, T., Liu, E., Chambers, N., 2020. Dream jobs? Teenagers' career aspirations and the future of work.
McKenzie, D., Mohpal, A., Yang, D., 2022. Aspirations and financial decisions: Experimental evidence from the Philippines. J. Dev. Econ. 156, 102846.
Ministry of Education, Youth and Sport, 2017. Public education statistics & indicators: 2016–2017. Ministry of Education, Youth and Sport. Department of Education Management Information System, Phnom Penh, Cambodia.
Ministry of Education, Youth and Sport, 2019. Public Education Statistics & Indicators: 2018–2019. Ministry of Education, Youth and Sport. Department of Education Management Information System, Phnom Penh, Cambodia.
Nguyen, T., 2008. Information, role models and perceived returns to education: Experimental evidence from madagascar. Job Market Paper.
OECD, 2025. The state of global teenage career preparation.
Papageorgiou, T., 2014. Learning your comparative advantages. Rev. Econ. Stud. 81 (3 (288)), 1263–1295.
Renée, L., 2025. The long-term effects of career guidance in high school and student financial aid: Evidence from a randomized experiment. Am. Econ. J.: Appl. Econ. 17 (2), 165–183.
Riley, E., 2024. Role models in movies: the impact of queen of katwe on students' educational attainment. Rev. Econ. Stat. 106 (2), 334–351.
Ross, P.H., 2019. Occupation aspirations, education investment, and cognitive outcomes: Evidence from Indian adolescents. World Dev. 123, 104613.
Serra, D., 2025. Role models in developing countries. In: Dasgupta, U., Maitra, P. (Eds.), Handbook of Experimental Development Economics. Edward Elgar Publishing, pp. 122–134.
UNESCO, 2011. Education and Fragility in Cambodia. International Institute for Educational Planning, Inter-Agency Network for Education in Emergencies, Paris, France.
UNESCO, 2020. Global Education Monitoring Report, 2020. United Nations Educational, Scientific and Cultural Organization, Paris, France.
United Nations Development Programme, 2023. Human Development Index (HDI) Data. (Accessed 03 September 2024).
van der Weide, R., Lakner, C., Mahler, D.G., Narayan, A., Gupta, R., 2024. Intergenerational mobility around the world: A new database. J. Dev. Econ. 166, 103167.
van Iddekinge, C.H., Roth, P.L., Putka, D.J., Lanivich, S.E., 2011. Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. J. Appl. Psychol. 96 (6), 1167–1194.