Journal of Development Economics

# Unlocking young women’s potential? The impact of a low-cost career guidance program

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | DEVEC-D-25-01517R1 |
| **Article Type:** | Registered Report Stage 2: Full Article |
| **Section/Category:** | Health, Education, gender, poverty |
| **Keywords:** | female labor force participation; career choices; gender norms; India |
| **Corresponding Author:** | Viola Asri<br>Chr Michelsen Institute<br>NORWAY |
| **First Author:** | Ankush Asri, PhD |
| **Order of Authors:** | Ankush Asri, PhD |
| | Viola Asri, PhD |
| | Anke Hoeffler, PhD |
| **Abstract:** | Societal expectations and gender norms constrain young people's career planning, particularly for girls, limiting the role of personal interests in skill-building decisions. This study evaluates a low-cost, 10-hour career exploration intervention through a pre-registered, school-level, clustered randomized controlled trial with over 6,000 primarily female students in urban India. The program improved future planning and increased the importance students placed on personal interests in career choices. However, there is no evidence of medium-term impacts on educational or skill-building investments. These findings suggest that while scalable interventions can shift aspirations and planning, structural barriers constrain longer-term behavioral change. |
| **Response to Reviewers:** | |

# Unlocking young women's potential? The impact of a low-cost career guidance program*

Ankush Asri[a,b], Viola Asri[a], Anke Hoeffler[b]

[a] *Chr. Michelsen Institute, Jekteviksbakken 31, Bergen, 5006, Vestland, Norway*
[b] *University of Konstanz , Universitätsstraße 10, Konstanz, 78464, Baden-Württemberg , Germany*

**Abstract**

Societal expectations and gender norms constrain young people's career planning, particularly for girls, limiting the role of personal interests in skill-building decisions. This study evaluates a low-cost, 10-hour career exploration intervention through a pre-registered, school-level, clustered randomized controlled trial with over 6,000 primarily female students in urban India. The program improved future planning and increased the importance students placed on personal interests in career choices. However, there is no evidence of medium-term impacts on educational or skill-building investments. These findings suggest that while scalable interventions can shift aspirations and planning, structural barriers constrain longer-term behavioral change.

*Keywords:* female labor force participation, career choices, gender norms, India
*JEL:* O10, O15, C93, I26, J13, J16

*This study was pre-registered under AEARCTR-0009977 with a detailed pre-analysis plan. We are grateful for inputs from Charity Troyer Moore, Rohini Pande, and participants at the North East Universities Development Consortium Conference 2024 and Advances with Field Experiments Conferences 2024. We thank our implementing partner organization, Alohomora Education Foundation, the local school authorities, and the Inclusion Economics India Centre at the Institute for Financial Management and Research (IFMR) for their collaboration and support. Sunidhi Agarwal, Ramanujan Ankam, Niharika Bhagavatula, Siddharth Jain, and Shubham Singh provided excellent research management and assistance. We gratefully acknowledge funding from the Excellence Cluster "The Politics of Inequality" at the University of Konstanz and the Alexander von Humboldt Foundation. Ethical approval was granted by the IFMR and the University of Konstanz.
*Corresponding author's email address:* viola.asri@cmi.no

## 1. Introduction

Low female labor force participation (FLFP) constrains economic growth, women's economic empowerment, and agency, which are essential for women and closely linked to the well-being of their households and their children (Duflo, 2012). India is one such country where, despite rapid urbanization and rising educational attainment, FLFP remains among the lowest globally, just 24% in 2022, and has been declining over the recent decades (International Labour Organization, 2022).[1]

Unlike in many other countries, where low FLFP is often linked to women exiting the labor force due to childbearing (Dex et al., 1998; Klasen, 2019), in India, most women never enter the labor force (NSSO, 2023). This is particularly striking in urban areas, where educational attainment is higher and more job opportunities are available. Yet, FLFP remains low, particularly for women who have completed secondary education but did not attain further skill formation (Fletcher et al., 2017).

Why do Indian women not participate in the labor force? Both demand-side and supply-side factors contribute to low FLFP. On the demand side, these include gender-based discrimination in hiring (Chowdhury et al., 2018; Riach and Rich, 2006; Carlsson, 2011; Kuhn and Shen, 2013), laws restricting women's access to work in certain industries and to work at night, as well as gender wage gaps (Deshpande et al., 2018; Duraisamy and Duraisamy, 2016). On the supply side, social norms around female mobility, household roles, and expectations shape women's decisions about post-secondary education and labor market entry (Jayachandran, 2020; Mehta and Sai, 2021). As a result, even women who desire to work often lack the necessary qualifications: more than half of non-working women who say they want a job report that they are under-qualified for the types of jobs they would prefer (Fletcher et al., 2017).

The career-mismatch is not just restricted to women; both young men and women face challenges in making career choices during the transition from secondary school to tertiary skill formation and are influenced by their

[1] More recent data shows a slight upward trend in FLFP but the debate on whether this due to changes in measurement or other more structural factors is ongoing and inconclusive (link).

parents (Ray et al., 2020). However, in contexts with rigid gender norms, the implications of suboptimal educational and career choices differ by gender. While men might pursue careers misaligned with their strengths and interests, women might not choose any career path, not complete skill formation, and not enter the labor force.

This raises an important question: Can better support during the transition from secondary school to further skill formation help young women pursue career paths that align with their interests and strengths? In particular, can early career guidance help address information gaps and low aspirations that may deter young people from making forward-looking educational investments?

Focusing on supply-side factors, we study the impact of a low-cost, school-based career exploration program (CEP) designed to support students during this critical transition period. The program aims to help students better understand their interests and strengths, and connect these to "feasible" career paths. Information on career options through targeted programs can mitigate market failures such as asymmetric information; however, they do not address the potential mismatch between students' interests and job requirements. Rather than simply delivering information or advice, the program uses interactive sessions and guided exercises to engage students in reflecting on their strengths, preferences, and opportunities. Further, the program is delivered by young professionals from similar backgrounds who can be seen as near-peer role models. We investigate whether the CEP can increase the share of students planning for a professional career and making relevant choices by taking entrance exams and enrolling in courses or training by addressing the information constraint, raising aspirations, and having role models. Our primary focus is on young women from disadvantaged socio-economic backgrounds. To minimize demand-side frictions like the availability of courses and jobs, we conduct the study in a metropolitan city where many career opportunities are accessible.[2]

To identify the causal impact of the career guidance program, we use a school-level clustered randomized controlled trial working with over 6000 students from 12th grade in 45 secondary non-fee paying schools. After collecting baseline data from May-July 2023 with self-administered surveys, the students in the treatment schools received the career guidance program from

[2]Based on our administrative approvals, we are not disclosing the name of the city.

July until November 2023, after which we collected the endline data again through self-administered surveys in both treatment and control schools. The endline survey data allow us to assess the program's causal impact on students' future plans and preparedness for career choices. Further, we collected follow-up data through a phone survey approximately 6 months after the students had completed school (12 months after the intervention) and had (mostly) made decisions on what they were doing. This enables us to assess the impact of the CEP on students' actual behavior related to career choices. Both the endline and the follow-up surveys were designed to capture potentially relevant mechanisms that can explain why and how the program impacts students' aspirations and choices. In addition, access to the attendance data collected by facilitators during program implementation allows us to examine how the program's impact differs for the students who participated in a higher number of sessions.

Our primary contribution lies in shedding light on how career guidance in school can facilitate students' career choices, focusing on young women from lower-income backgrounds in a setting with rigid gender norms. Existing causal evidence on the impact of career guidance stems primarily from developed countries and focuses on one selected choice, such as choosing the academic high school track (Carlana et al., 2022), enrolling in college (Bettinger and Evans, 2019), or selecting a major (Baker et al., 2018). We instead focus on the choice to continue skill formation, accounting for different ways of doing so, and assess the impact of a relatively low-cost CEP in which, in contrast to career counseling, students play an active role but are guided by young facilitators while exploring career options corresponding to their interests and talents.

Our findings highlight both the potential and the limitations of interest-based career guidance in shaping young people's transitions beyond school. On the one hand, we find that the CEP significantly improved students' plans for their transition to further skill formation. Students in treatment schools were more likely to plan to take entrance exams for tertiary education and consider their interests the most important factor in choosing a career path. Our results also suggest that the CEP increased the share of students believing in intergenerational mobility. The estimates of the impacts double for students who participated in sessions regularly, underscoring the importance of institutional support provided by schools for implementing such a program. Further, mediation analysis reveals that the observed impacts are primarily driven by the additional information provided by the program and

the facilitators serving as role models.

On the other hand, these positive shifts in students' career plans did not translate into a behavioral change. In the follow-up survey, we do not find significant differences between treatment and control students about one year after the intervention (six months after finishing secondary school). Students from treatment schools are not more likely to be enrolled in further education or training, nor are they more likely to have taken an entrance exam. While we do not have detailed time-use data, our exploratory analysis shows that the students from the treatment group report spending less time helping in the household compared to the control group, which may suggest that they are more focused on their careers. However, this shift did not materialize into measurable education or career investments, at least within the observed time frame. Our heterogeneity analysis shows that short-term impacts are stronger for students from households where no woman has been working for the last two years, students who are more mobile, and students who strongly believe in meritocracy. Impacts beyond the short term only last for a few subgroups, such as co-educational students being more likely to be enrolled in a course and non-Hindu students to be working due to the CEP.

We interpret these findings as reflecting that structured guidance may shift how students think and plan, which is also in line with the previous literature. However, material and social constraints may continue to limit what students do. Especially for young women, the conversion of aspirations into action may be hindered by household financial pressure, parental expectations, societal pressures, or logistical barriers. This highlights how context shapes both the process and the limits of aspirational change.

The remainder is structured as follows: Section 2 describes the existing literature and our contribution, Section 3 provides the contextual background information and explains the experimental design, and Section 4 presents the data and the empirical strategy. Section 5 describes our results, and section 6 discusses potential policy implications and avenues for future research.

## 2. Literature

With this paper, we contribute to four strands of literature: First, we add to the growing body of work on aspirations shaped by individual experience and social context. As Genicot and Ray (2020) argue, preferences are formed within the societal structures in which individuals live. The concept

of the "cognitive neighborhood" underscores how social environments influence individuals' goals, shaping their future choices. While aspirations can be a powerful driver of personal development, overly ambitious or unrealistic goals may lead to disillusionment and frustration. The program we evaluate takes this into account by encouraging students to form realistic aspirations and set achievable objectives, while explicitly considering financial and contextual constraints. It also promotes the development of contingency plans in case the primary goal proves unattainable or overly risky.

Second, our study contributes to the literature on students' perceptions of the returns to education and skill formation, and how information interventions can improve educational outcomes, both in terms of continuing education and selecting its type. Jensen (2010) shows that eighth-grade boys in the Dominican Republic significantly underestimate the returns to secondary education, and that providing school-level data on actual returns raises educational attainment. In a related study, Jensen (2012) finds that introducing recruitment services for newly available call center jobs in Indian villages led to increased educational investment, delayed marriage and childbearing, and higher female labor force participation. Going beyond expected earnings, Wiswall and Zafar (2015) emphasize the role of perceived ability and heterogeneous preferences, showing that a combination of anticipated returns, self-assessed competence, and intrinsic interests influences career decisions. We contribute to this literature by evaluating a program that not only provides information about the returns to education and skill development, but also explicitly encourages reflection on personal interests and individual strengths.

Third, our research contributes to the literature on career guidance and mentoring interventions, which has largely concentrated on outcomes such as academic track selection, college enrollment, participation in STEM studies, and labor market orientation, primarily in high-income countries including Italy, the United States, Canada, and Germany. Carlana et al. (2022) demonstrate that providing tutoring to high-achieving immigrant students in Italy increases the likelihood of male students enrolling in academically oriented high school tracks. Bettinger and Evans (2019) find that assistance from recent college graduates during the application and enrollment process significantly boosts college enrollment among Hispanic and low-income students. Porter and Serra (2020) show that exposure to successful, charismatic female Economics graduates from the same university encourages increased enrollment in Economics courses among female students. Renée (2025) inves-

tigates the long-term effects of career guidance versus financial aid in Canadian schools. While both interventions improved college enrollment rates, the career guidance program yielded more persistent impacts on long-term earnings. Resnjanskij et al. (2024) examine a mentoring initiative and find sustained positive effects on students' grades and labor market orientation, particularly among those from low socioeconomic backgrounds.

While our study also employs relatable facilitators—what Bettinger and Evans (2019) term near-peers—we extend beyond assessing single educational decisions (e.g., track selection or college enrollment) to investigate more comprehensive support for continued skill formation. Informed by evidence from high-income contexts, our intervention is implemented in a resource-constrained environment and targets students from low-income households, leveraging the type of school in which the program is embedded. This design is informed by findings that interventions typically yield greater effects for more disadvantaged students (Bettinger and Evans, 2019; Carlana et al., 2022), and that career guidance is most effective for individuals least likely to access or seek it independently (Zafar, 2019; Lee and Ekstrom, 1987).

Fourth, we contribute to the limited body of evidence on career-related interventions in developing countries, which has predominantly focused on very short-term (often one-off), low-cost approaches aimed at providing information on educational returns and career planning skills to high school students. Existing studies from China and Cambodia report mixed or even adverse effects. Loyalka et al. (2013) show that offering information on returns to education and career planning skills in Chinese junior high schools did not improve educational outcomes; in fact, counseling increased dropout rates and lowered academic achievement. Similarly, Gehrke et al. (2023) find that a half-day career planning workshop raised high school enrollment and attainment among high-performing students but had adverse short-term effects on low-performing students. Kipchumba et al. (2024) investigate whether visits from university students acting as role models can influence aspirations and gender attitudes among high school students. They find that six months after the intervention, exposure to female role models improved students' attitudes toward gender equality but had no measurable effect on aspirations to attend college.

Our study is also situated in a developing country and features a low-cost, scalable intervention. However, it departs substantially from traditional counseling approaches in which counselors prescribe what they believe is best for the student. In this NGO program, students remain the central decision-

makers, supported by facilitators who guide them in exploring career options and opportunities, both collaboratively in small peer groups and individually. The design and timing of the study allow us to assess impacts during a critical transition period, with baseline data collected at the start of 12th grade, endline data at the conclusion of 12th grade, and a follow-up approximately six months after school completion.

In summary, this study contributes to the literature by emphasizing career exploration over mere information provision or high-cost counseling. The program actively involves all students, encourages the consideration of diverse pathways, and integrates personal interests, individual strengths, and the economic returns to education. Its low cost and scalability make it particularly well-suited to resource-constrained contexts, offering the potential for broad and sustained impact.

## 3. Background and experimental design

### *3.1. Context*

To mitigate potential demand-side constraints, we designed this study in the context of a large urban labor market that offers a wide range of opportunities for skill development and employment. Accordingly, we implement the randomized controlled trial (RCT) in schools in a metropolitan area in Northern India. These are no-fee schools that primarily serve students from low- to lower-middle-income households in the surrounding neighborhoods. Consistent with this context, we observe that many students in our sample—or their older siblings—are the most educated members of their families. Their parents typically have not completed secondary education, and mothers tend to have even lower educational attainment than fathers. Furthermore, a significant share of students report never receiving career guidance, including informal advice. A more detailed overview of the students' baseline characteristics is provided in Section 5.1, which further contextualizes the sample for our analysis.

Data from representative surveys in this metropolitan context highlight that not being in education, employment, or training (NEET) is a substantial challenge, particularly among young women aged 18–30. According to the National Family Health Survey (International Institute for Population Sciences and ICF, 2022), while 66 percent of young women who complete secondary education enroll in tertiary education, only 20 percent subsequently enter the labor force. More broadly, data from the National Sample Survey

indicate that 64 percent of young women with a secondary school education fall into the NEET category. Among those pursuing post-secondary education, college enrollment is the predominant route, whereas vocational training remains rare; only 8 percent of women in this setting opt for vocational courses. Despite evidence that vocational education strongly predicts future employment outcomes, this corresponds with many women's expressed willingness to work (National Sample Survey Organization, 2021).

The intervention studied in this paper seeks to reduce the incidence of NEET status among young people by encouraging interest-based career decision-making, expanding students' consideration of diverse tertiary education pathways, and supporting the development of actionable, individualized plans to achieve their career goals.

### *3.2. Intervention*

We collaborate with the relevant school authority to provide school access and with a local NGO to implement the intervention. Official approval to conduct the project in 45 secondary schools was obtained through the research initiative Inclusion Economics India Centre at the Institute of Financial Management and Research. Our partner NGO, Alohomora Education Foundation, originally designed the intervention in 2014–2015 and has since then been implementing and refining the program, primarily working with 11th- and 12th-grade students in schools located in low-income urban neighborhoods similar to those included in the RCT.

The Career Exploration Program (CEP) consists of 15 sessions, each lasting 40–45 minutes, and is delivered in classrooms during regular school hours. The program is designed to guide and encourage students to recognize their strengths, interests, and talents; reflect on personal barriers and constraints; explore and identify suitable career options; seek out potential sources of support; and develop actionable plans to pursue a chosen career path.

The program follows a five-step process that helps students make career choices aligned with their interests and abilities. The NGO provides the following description of each step.[3]

1. **Start to think** - "Thinking early about making a choice lets one take charge of their career pathway, giving time at hand to explore."

[3]The original text is in Hindi and has been translated by the authors.

2. **Think openly** - "Not limiting yourself to a few conventional careers, but thinking openly, for, e.g., thinking of the healthcare sector instead of just being a doctor, opens one up to a variety of options to choose from."

3. **Understand yourself** - "Experiencing different types of work, getting feedback, and reflecting on those experiences to find which careers align more with yourself."

4. **Prepare to succeed** - "Creating a plan of action to reach the chosen goal, including the research of where to go and how, and steps to prepare oneself for the same."

5. **Thrive in your chosen path** - "Taking continuous action-reflection cycles to up-skill and succeed in the work."

Throughout the five stages, the program incorporates the use of a virtual portal designed to offer students a comprehensive learning experience about various careers. Early in the program, each student creates a personal, password-protected profile on the portal. The program concludes with a "showcase" event, where all participating students present their career objectives and plans of action to an audience of 12th-grade peers, teachers, and program facilitators.

The program is designed to help students set realistic expectations by taking into account constraints such as financial limitations, academic ability, and regulations specific to the Indian education system.[4] Facilitators guide students through a structured reflection on their personal circumstances and financial constraints. They are trained to be transparent about the competitiveness of certain career paths, such as becoming a civil servant, and to encourage students to develop backup plans when pursuing high-risk or highly competitive options.

[4]For example, university admission for several subjects requires students to have chosen a corresponding stream in secondary school. Students can enroll in Science, Commerce, or Arts & Humanities streams, but not all no-fee schools offer the Science stream. The Science stream is the most prestigious, providing access to fields such as medicine or engineering. The Commerce stream follows in reputation and is required for subjects like Economics or Business Administration. Stream selection is strongly correlated with students' grades but may also be influenced by their parents.

In each classroom and during each session, two facilitators are engaged to implement the program, with one facilitator leading and the other supporting activities. Facilitators are gender-matched with the students (i.e., female facilitators in girls' schools, male facilitators in boys' schools, and mixed-gender pairs in co-educational schools). They are, on average, four years older than the participating students, typically enrolled in tertiary education, and come from similar communities—factors that help students relate to them. In addition to in-person sessions, facilitators use WhatsApp groups to interact with students and foster peer-to-peer communication. They also conduct one-on-one conversations to address individual questions and challenges.[5] We provide more detailed information about the program in Appendix B.

### *3.3. Causal pathway*

Figure 1 presents our theory of change for the Career Exploration Program (CEP) and its expected impact on students' career intentions and decisions. The program targets market failures like imperfect information, low aspirations, and lack of role models—factors that drive inefficient human capital allocation and suboptimal labor supply choices.

The top row illustrates the underlying **problems** as observed during our qualitative and quantitative formative work: students have limited self-awareness about their interests and strengths, leading to poorly aligned career paths. This is particularly acute in low-cost schools lacking formal guidance, where socioeconomically disadvantaged students rely on inaccurate or incomplete information. Many face pressure to enter the informal labor market early due to financial constraints and a limited understanding of education's long-term benefits.

External influences, such as family and societal expectations, frequently steer students toward a narrow range of "widely respected" professions like teaching or civil service, while undervaluing other options. Gender norms further restrict girls' aspirations, often preparing them for domestic roles rather than independent careers. The absence of relatable role models reinforces these dynamics and suppresses ambition.

As documented in our formative qualitative and quantitative fieldwork with female students, around 70% of students tend to be more educated than their parents, and about half report never receiving any form of ca-

[5]The program implementation costs about 30-35 USD per student.

reer guidance, not even informal advice from family members or relatives. Students face a lack of professional role models: nearly 60% report that no woman in their household has worked outside the home in the past two years. Despite these students' relatively higher educational attainment, traditional gender norms remain important, with 31% believing in a woman's primary role to be the caretaker of the household, and 28% agreeing that men should be prioritized when jobs are scarce. In this context, becoming a teacher or a civil servant is the most common career aspiration.

The CEP addresses these challenges through various **inputs**, shown in the second row. Facilitators lead classroom-based sessions during school hours that help students explore personal interests, strengths, and career options. Activities include group discussions, planning exercises, and support in developing and presenting individual career action plans. Students are also guided to communicate their plans with parents and identify potential mentors or supporters.

The intended **outputs** include greater student awareness of their interests and constraints, improved understanding of education's value, and increased knowledge of diverse careers. Students are expected to engage in peer discussions, view facilitators as mentors, and share career plans with parents or other supporters.

Observing these outputs relies on the following **assumptions**: that students are curious about career options and education, open to sharing plans with peers, able to relate to facilitators, and capable of discussing aspirations with adults around them. Importantly, we also assume that students have some degree of agency in shaping their educational and career paths.

From these outputs, we expect short- and medium-term **outcomes**. In the short term, students may plan to enroll in training, prepare for entrance exams, and aim for employment. They may also become more optimistic about intergenerational mobility. In the medium term, we anticipate observable behavioral changes—actual enrollment in courses, exam participation, or completion of short skill-building programs like English or computer classes.

Our causal chain rests on five potential **mechanisms**, which we will examine empirically:

1. The CEP provides missing information about career options.
2. The CEP raises aspirations for education and expected earnings.
3. The CEP encourages peer discussions around career planning.

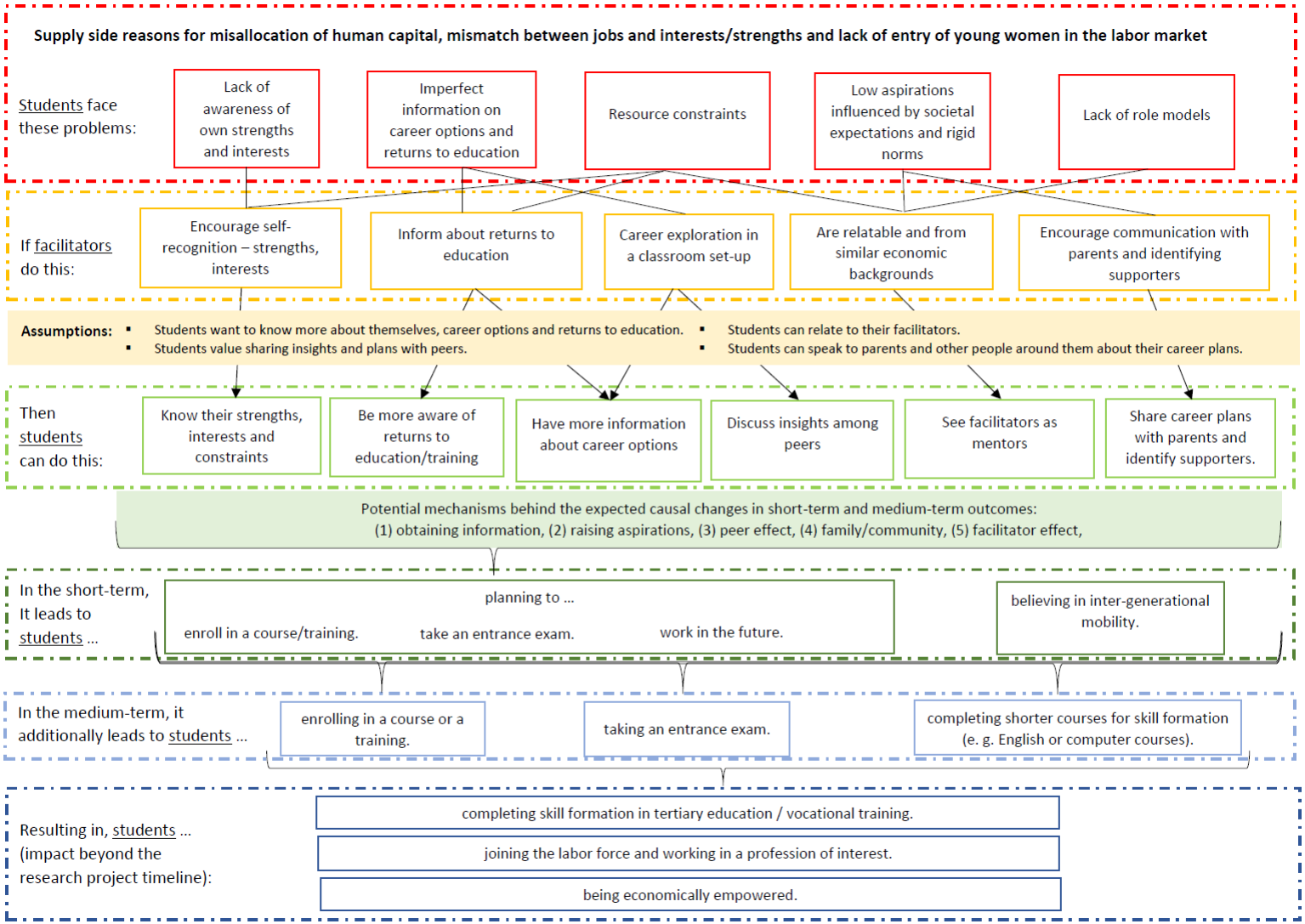
4. The CEP helps students gain support from parents and others.

5. The CEP facilitators serve as mentors and role models.

Given the program's potential to expand career opportunities, especially for young women, we describe our empirical strategy to test these mechanisms in Section 4.3.

If these medium-term outcomes occur, long-term **impacts** after several years, and thereby beyond the research project timeline, could be observed too, including the CEP supporting students to complete their skill formation, to join the labor force, to work in a profession of interest, and to be economically more empowered as shown in Figure 1.

Figure 1: Theory of change



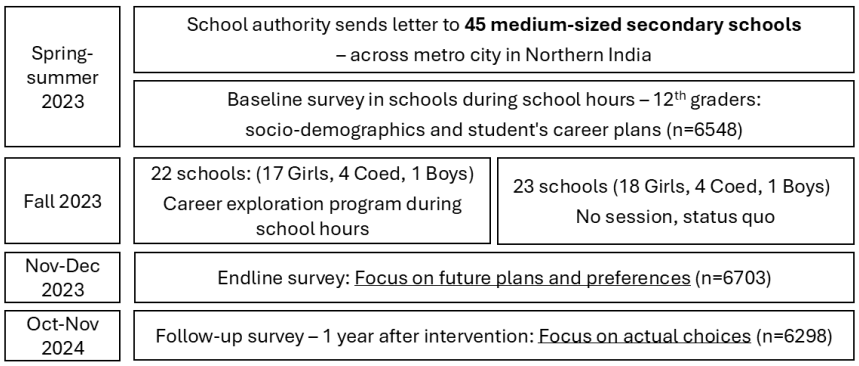
### *3.4. Experimental design and randomization*

Our collaboration with the relevant school authority and partner NGO enables us to conduct the school-level clustered RCT with 45 schools. The randomization is stratified by school type (girls, boys, and co-educational schools), with 22 schools in the treatment group and 23 schools in the control group. The flowchart in Figure 2 visualizes our procedures.

Starting in Spring 2023, the school authority informed all 45 senior secondary schools about the data collection procedures and treatment schools' requirements for the program implementation in terms of required rooms and time. Subsequently, in Summer 2023, research associates supported by small teams of surveyors conducted the baseline surveys in all schools, focused on students' socio-demographics and career plans. The CEP was implemented in 22 schools during school hours, and no intervention was implemented in the remaining 23 schools. The endline surveys focused on students' career plans were conducted in November-December 2023, and the follow-up data collection was completed as a phone survey approximately one year after the intervention implementation in October-November 2024. Appendix A provides the project timeline and crucial events from a student's perspective.

We chose school-level randomization to minimize the probability and extent of spill-over effects (Muralidharan, 2017). School principals, teachers, and students were not informed about the detailed study objectives and the research-specific assignments in the treatment and control groups. We implement this by informing school principals and teachers only about what will happen in their school and our broader motivation for data collection to assess the need for career guidance in low-cost schools. No information is provided on what is happening in other schools. Given that school-specific activities happen frequently in this context and are often organized by NGOs or the school authority for selected schools, we expect any spill-over effect to be rather unlikely and minor, if at all, to lead to an attenuation bias for the measured impact.

Figure 2: Research procedures




Source: Authors' visualization.

## 4. Empirical strategy

We submitted our study as a Pre-Results Review Paper to the *Journal of Development Economics*, which accepted it on September 7, 2023, prior to the endline surveys. The registered report is available as an Online Appendix to this paper. We follow the pre-registered analysis throughout and explicitly indicate where we present additional exploratory analyses. The field experiment was implemented as planned in the registered report.

### *4.1. Data and hypotheses*

Our analysis relies on three main data sources: (1) primary data from baseline, endline, and follow-up surveys that we conducted; (2) administrative data provided by schools; and (3) data from the partner NGO regarding the sessions. The baseline and endline surveys were completed by the students themselves on tablets provided in the classroom.[6] The response rate in

[6] As pre-registered, we used self-administered surveys to reduce costs, facilitate data collection with more than 6000 students in classroom settings, and likely limited interviewer-related demand effects by minimizing direct social interactions. At the same time, self-administered surveys have potential limitations. Students may have been less attentive, particularly those in control schools who did not participate in the intervention, and the absence of surveyors may have altered reporting incentives in either direction. The results stay the same when controlling for the attentiveness of the students (see section Appendix E), although only 129 students were inattentive in the endline surveys (67 in control and 62 in treatment), which have the primary outcome variables.

each of these surveys was approximately 75 percent, consistent with regular attendance rates in this school setting. The follow-up survey on students' actual choices was conducted by phone, using the contact numbers collected during the baseline and endline surveys. As we required at least one phone number per student, we contacted all students who participated in at least one of the two surveys. The follow-up survey achieved a similarly high response rate, with data collected from approximately 80% of the students.

The baseline survey (n=6548) includes socioeconomic information on the students, a module measuring social desirability, and data on students' access to career guidance both privately and through their schools. The endline survey (n=6703) focuses on students' career plans and potential mechanisms such as increased information, aspirations, and social support.[7] While the endline survey captures students' intended career choices while still in school, the follow-up survey (n=6298) measures their actual choices after leaving school.

Building on the theory of change in Section 3.3, we pre-registered the following hypotheses:[8]

In the short term, based on the causal pathways outlined in our theory of change, we expect that the CEP increases the probability that a student plans to enroll in a course or training (H1), plans to take an entrance exam (H2), plans to work outside the home for pay (H3), selects a profession based on interests (H4), and believes in intergenerational mobility (H5).

In the longer term, we expect the CEP to increase the likelihood that a student has enrolled in a skill-development course (H6), has taken a competitive exam for admission or employment in certain sectors (H7), has completed a skilling program (for shorter courses) after finishing 12th (H8) in case not currently enrolled in a course or training, and has started working outside the home for pay (H9).

[7] While the greater sample size in the endline is directly linked to students' school attendance, our survey teams also noticed that many students enjoyed taking the surveys, which may also explain why endline participation rates were even higher than baseline participation rates.

[8] In our pre-results review and pre-analysis plan, we distinguished between "primary" and "secondary" outcomes. In the present paper, we present outcomes chronologically (short-term, medium-term, and exploratory), without labeling them as primary or secondary. This choice was made for clarity of exposition and ease of interpretation. The substantive set of outcomes and hypotheses remains unchanged from the pre-analysis plan.

We also examine potential mechanisms through which these impacts may occur. These include increased access to information due to the CEP, elevated educational or financial aspirations, and enhanced social support from the broader community and directly from the CEP facilitators, who act as role models.

Finally, we test pre-registered secondary hypotheses regarding the effects of the CEP on locus of control and self-efficacy, both in the short term (H10–H11) and the long term (H12–H13).

*4.2. Regression analysis*

All our estimations follow the pre-analysis plan unless explicitly mentioned as exploratory. We begin by conducting balance checks to compare baseline characteristics across individuals in treatment and control schools. We estimate the following regression:

$$X_{0isd} = \alpha + \beta T_s + \gamma_d + \epsilon_{is}, \tag{1}$$

where $X_{0isd}$ denotes a baseline covariate measured at time $t = 0$ for individual $i$ in school $s$ and district $d$, $T_s$ is a binary indicator equal to one if school $s$ was assigned to the treatment group, and $\gamma_d$ represents district fixed effects. The covariates include age, gender, the stream they are in (arts, commerce, or science), if they have a sibling, having an older brother or an older sister, if they have previously received career counseling, household size, the total assets owned by the household, and the social desirability index from baseline surveys. Standard errors are clustered at the school level. Since treatment is assigned at the school level and was stratified by school type and district, we include school type and district fixed effects to account for stratification. Each regression also controls for the baseline value of what the students had reported in the baseline, if available. That is, while testing hypotheses H1 and H6, we also control whether the student reported planning to enroll in a course or training during the Baseline surveys. Similarly, for hypotheses H2 and H7, if the student reported planning to take an entrance exam, and for H3 and H9, if the student had reported planning to work in the future for pay during the baseline surveys.

Next, we estimate the intention-to-treat (ITT) effect of the intervention on outcomes of interest. Our main estimating equation is:

$$Y_{is} = \alpha_1 + \beta_1 T_s + \delta X_{0is} + \gamma_d + \omega_{is}, \tag{2}$$

where $Y_{is}$ denotes the outcome for individual $i$ in school $s$, and $X_{0is}$ includes a set of baseline covariates. The coefficient $\beta_1$ captures the ITT effect. Standard errors are again clustered at the school level.

In addition to estimating the ITT as pre-registered, we estimate the intervention's local average treatment effect (LATE) on students who regularly participated in the program. We define regular participation as completing at least 10 sessions, corresponding to the median number of sessions attended in the treatment group. Since participation is potentially endogenous, we instrument regular participation using the randomized treatment assignment at the school level. We implement a two-stage least squares (2SLS) approach.

The first-stage regression is specified as:

$$D_{is} = \pi_0 + \pi_1 T_s + \gamma_d + \eta_{is}, \tag{3}$$

where $D_{is}$ is a binary indicator equal to one if individual $i$ participated regularly in the program.

The second-stage regression, which yields the LATE estimate, is:

$$Y_{is} = \alpha_2 + \beta_2 \hat{D}_{is} + \delta X_{0is} + \gamma_d + \nu_{is}, \tag{4}$$

where $\hat{D}_{is}$ is the predicted value of regular participation from the first stage. The coefficient $\beta_2$ captures the causal effect of the intervention on compliers, i.e., those whose participation is influenced by treatment assignment. As before, we cluster standard errors at the school level. We correct the p-values for multiple-hypothesis testing using the approach by Anderson (2008) and mention the obtained q-value in the text.

In all specifications, the sample size depends on whether the analysis focuses on short-term outcomes using data from the endline and baseline or medium-term outcomes using data from the follow-up surveys and baseline.

### *4.3. Mediation analysis*

We see alleviating information constraints as the primary mechanism through which the CEP is expected to affect students' decisions. However, other channels may also contribute to its impact in addressing the potential misallocation of human capital in general, especially among young women who risk making career choices that do not match their interests and may not enter the labor market at all. To explore these mechanisms, we conduct a mediation analysis following the framework developed by Heckman et al.

(2013) and Heckman and Pinto (2015).[9]

This mediation framework allows us to quantify how much of the program's total impact on key outcomes can be attributed to specific observable channels, or "mediators". We focus on six such mechanisms: (1) improved access to information, (2) raised educational/financial aspirations, (3) peer effects, (4) family and community support, and (5) facilitator influence. The analysis decomposes the overall treatment effect into six components, five attributed to these mediators and one residual component capturing unexplained variation.

Assuming that the outcome can be expressed as a linear combination of the $k = 5$ mediators $M_i^k$ and a vector of baseline demographic characteristics $X_0is$, we estimate the following model:

$$Y_{is} = \alpha + \beta^{\text{residual}} T_{cs} + \sum_k \theta^k M_i^k + \delta X_{0is} + \gamma_s + \varepsilon_{is} \tag{5}$$

Here, $\beta^{\text{residual}}$ captures the portion of the treatment effect not explained by the observed mediators. The share of the total effect explained by the mediators is given by $1 - \frac{\beta^{\text{residual}}}{\beta}$, where $\beta$ is the total treatment effect.

To estimate the impact of the treatment on each mediator, we use:

$$M_{is}^k = \alpha_o^k + \beta_1^k T_{is} + \delta_1^k X_{0is} + \gamma_{0_s} + v_{is} \tag{6}$$

The contribution of each mediator $k$ to the overall effect is then calculated as:

$$ShareM_k = \theta \frac{\beta_1^k}{\beta_1} \tag{7}$$

Therefore, the kth mediator's contribution to the share of the overall treatment effect is the direct effect of the mediator, $\theta^k$, weighted by the proportion of the treatment effect on the mediator and the total treatment effect.[10]

This decomposition helps us to assess the extent to which the program's effect on students' decisions is through the five mechanisms. Each of these

[9] For an application of this approach, see Resnjanskij et al. (2024).

[10] This estimated effect is an upper bound of the mediator effect since the mediation effect relies on the assumption that the error term, including any unobserved mediators, is orthogonal to the included mediators.

mediators addresses a different constraint in the process of skill development and labor market entry:

**Access to information**: The program introduces students to a variety of career options, helps them understand the value of skill formation, and outlines the steps needed to reach their goals. This helps increase the likelihood of continuing skill formation and diversifying their career options while considering financial constraints. We ask students to write down up to three occupations that they think would be good for them, and for each career they enter, we ask them how well informed they are about the career. Evidence for this channel would suggest that the CEP improves decision-making by reducing informational frictions.

**Raising aspirations**: By engaging with facilitators, peers just a few years ahead, and discussing futures with classmates, students may develop higher aspirations. We measure this through students' educational aspirations and salary expectations for their first job.

**Peer effects**: Given the program implementation in the classroom, students are encouraged to discuss their career plans with their peers and explore different career options jointly. These interactions among students increase the likelihood of students attending sessions, exploring different careers, and reflecting on them. It can also foster teamwork and fill the gaps in their knowledge, as students are expected to share complementary information with each other. We capture this potential peer effect in two ways.

- Peer communication: This channel captures whether talking to peers is a relevant source of information for students' future choices or whether talking about the future is a topic of discussion among students' peers.
- Peer influence: This channel captures the extent to which students are influenced by their friends, planning to continue skill formation in the endline or continuing skill formation in the follow-up, using the data from up to six other students in their grade that they speak to the most.

Evidence for this mechanism would imply that the CEP affects students' career choices through peer effects and that students in their

networks can more effectively process and use information than individual students.

**Family/community support**: The CEP encourages students to share their career objectives with family members and/or people in their community to identify potential supporters. This addresses the potential lack of social support in pursuing a career objective. We measure this by asking students whether they get any relevant information from a family member or someone in their community, and whether they get support from a family member or someone in their community to achieve their professional objectives. Evidence for this mechanism would imply that the CEP improves students' career choices by encouraging them to obtain support from their families and or people in their community.

**Facilitator effects**: The facilitators leading the CEP sessions are from similar backgrounds and schools, but have already made career decisions. They may act as role models, helping students update their beliefs about their ability to pursue certain career paths. We ask students whether there is someone around them who inspires them, whom they know personally, with one of the answer options referring to the CEP facilitator. Evidence for this mechanism would imply that the CEP can improve students' career choices by providing a role model similar to role model interventions in other settings (Lafortune et al., 2018; Porter and Serra, 2020; Kipchumba et al., 2024).

### *4.4. Heterogeneity*

We examine heterogeneous treatment effects among subgroups of girls who may face different constraints due to prevailing social norms using pre-treatment variables collected during the baseline. Specifically, we focus on analyzing variation in program impact along three dimensions - gender norms at home, mobility, and beliefs in meritocracy:

(i) **Patriarchal gender norms:** Students from less patriarchal households may benefit more from the CEP than those from more patriarchal backgrounds. We define households as non-patriarchal if the student reports that a woman in the household has worked for pay within the past two years. In our sample, 39% of female students live in households where a woman has worked (non-patriarchal), while 61% live in households where no woman has worked (patriarchal).

(ii) **Mobility:** More mobile students, that is, those permitted to travel independently to more places, may experience larger benefits from the CEP. We assess students' mobility by asking whether they are allowed to go alone to the local market, outside their neighborhood, or to a friend's home. We categorize students as "more mobile" if they are permitted to visit at least two of these three places (36% of the sample), and "less mobile" if they are allowed to visit only one or none (64%). Specifically, 16% of students are allowed to go to all three places, 20% to two places, 26% to one place, and 38% to none.

(iii) **Meritocratic beliefs:** Students who hold stronger meritocratic beliefs and agree more strongly that success depends on individual effort may gain more from the program. While we planned a median split here, beliefs in meritocracy show limited variation: 68% of students select the highest possible rating (10) on a scale assessing the belief that hard work pays off. We classify students as holding strong meritocratic beliefs if they rate it as 10 and as not holding such beliefs otherwise.

In addition to these, we also consider other sources of heterogeneity that seem to be relevant in this context: Arts vs. non-arts stream, female vs. male students, co-educational vs. single sex schools, upper caste vs. non-upper-caste, Hindu vs. non-Hindu, and academically high-performing vs. non-high-performing according to 10th-grade board exam grades.[11]

## 5. Results

### *5.1. Balance checks*

As shown in the balance checks in Table 1, nearly all variables are balanced between the treatment and control groups, except age. On average, students in the control group are 0.12 years older than those in the treatment group at baseline.[12] We control for this age difference in all specifications.

[11] To respect local regulations regarding data collection during school hours in schools, data on caste and religion could only be collected during the follow-up surveys, which were conducted over the phone.

[12] This difference is mechanical, since the baseline survey in control schools was conducted approximately six weeks later than in treatment schools, due to logistical constraints.

Table 1: Balance table at the student level from student surveys

| Variable | (1) Control | (2) Treatment | (3) Diff. (T-C) | (4) Obs |
|---|---|---|---|---|
| Female | 0.898 | 0.883 | -0.015 | 5,381 |
| | (0.303) | (0.322) | (0.072) | |
| Age | 16.836 | 16.714 | -0.122*** | 5,381 |
| | (0.753) | (0.751) | (0.040) | |
| Older brother | 0.462 | 0.434 | -0.028 | 5,381 |
| | (0.499) | (0.496) | (0.021) | |
| Older sister | 0.452 | 0.438 | -0.014 | 5,381 |
| | (0.498) | (0.496) | (0.021) | |
| Self most educated | 0.318 | 0.323 | 0.005 | 5,376 |
| | (0.466) | (0.468) | (0.017) | |
| Received guidance before | 0.518 | 0.538 | 0.020 | 5,381 |
| | (0.500) | (0.499) | (0.034) | |
| Uses smartphone | 0.927 | 0.911 | -0.016 | 5,381 |
| | (0.261) | (0.285) | (0.012) | |
| Owns smartphone | 0.214 | 0.203 | -0.011 | 5,381 |
| | (0.410) | (0.402) | (0.019) | |
| Risk preference | 7.031 | 7.130 | 0.099 | 5,381 |
| | (2.799) | (2.704) | (0.142) | |
| Patience | 7.575 | 7.619 | 0.045 | 5,381 |
| | (2.776) | (2.715) | (0.122) | |
| Social desirability index (1-6) | 4.402 | 4.346 | -0.056 | 5,381 |
| | (1.106) | (1.135) | (0.039) | |
| Belief in meritocracy (1-10) | 8.576 | 8.576 | -0.000 | 5,381 |
| | (2.537) | (2.471) | (0.082) | |
| Women more than housewives | 0.623 | 0.624 | 0.001 | 5,381 |
| | (0.485) | (0.484) | (0.026) | |
| Women access to jobs | 0.696 | 0.703 | 0.008 | 5,381 |
| | (0.460) | (0.457) | (0.021) | |
| Observations | 2,958 | 2,423 | 5,381 | |

Standard errors are clustered at the school-level and shown in parentheses. Baseline survey 2023.

The sample is primarily female, with 89% of students identifying as girls, and the average student is just under 17 years old. Most students report having at least one older sibling, and 32% are the most educated individuals in their households. About 52% of the students report receiving some guidance on their careers, including an informal one, before the baseline. While 92% of students use a smartphone, only 21% own one. The students score relatively high on measures of risk tolerance (7.0 out of 10) and patience (7.5 out of 10). Their responses show a tendency to give socially desirable answers (4.4 out of 6) and a strong belief in meritocracy (8.6 out of 10), with most students selecting 10 standing for "hard work usually pays off". Attitudes toward gender roles seem relatively progressive: 62% agree that women's roles go beyond domestic responsibilities, and 70% think women should have the same access to jobs as men when jobs are scarce.

Table 2 shows that students come from households with an average size of 5.4 members and report ownership of 3 out of 6 household assets (i.e., motorcycle/scooter, car, television, washing machine, fridge, and air conditioner). Female-headed households make up 16% of the sample, but only 39% of students report that any woman in their household has worked in the past two years. Parental education levels are relatively low, with mothers having completed an average of 6.4 years of education and fathers 8.7 years. Several students, however, did not know their parents' education levels.

Administrative data in Table 3 show that about 70% of students are enrolled in the arts and humanities stream, 20% in commerce, and only 9.5% in science. However, the low percentages of students being enrolled in science and commerce are likely due to the type of schools we are working with, offering commerce streams only to a minority of well-performing students and often lacking a science teacher and thereby a science stream.[13] Single-sex schools are common among the non-fee paying schools in this urban setting, and with our focus on young women, 78% of our sample attend girls' schools, 18% attend co-educational schools, and only 4% attend boys' schools. On average, there are 170-189 students in a 12th-grade class of a school and 40 students in a section, with a school having on average 4.9 sections in the 12th grade.

[13]For example, Nouman (2022) describes this problem.

Table 2: Balance table at the household level from student surveys

| Variable | (1) Control | (2) Treatment | (3) Diff. (T-C) | (4) Obs |
|---|---|---|---|---|
| Household size | 5.485 | 5.414 | -0.070 | 5,381 |
| | (1.692) | (1.782) | (0.093) | |
| Nr. of assets out of 6 | 3.009 | 3.073 | 0.065 | 5,381 |
| | (1.471) | (1.543) | (0.128) | |
| Female head of household | 0.159 | 0.163 | 0.004 | 5,381 |
| | (0.365) | (0.369) | (0.011) | |
| Any woman works | 0.380 | 0.402 | 0.023 | 5,381 |
| | (0.485) | (0.490) | (0.023) | |
| Mother's education in years | 6.292 | 6.559 | 0.268 | 5,102 |
| | (4.639) | (4.620) | (0.409) | |
| Father's education in years | 8.652 | 8.788 | 0.136 | 5,055 |
| | (4.137) | (4.000) | (0.338) | |
| Observations | 2,958 | 2,423 | 5,381 | |

Standard errors are clustered at the school-level and shown in parentheses. Students reported on their parents' education, and some did not know their parents' education levels, leading to missing values. Baseline survey 2023.

Table 3: Balance table at the school level using administrative data

| Variable | (1) Control | (2) Treatment | (3) Diff. (T-C) | (4) Obs |
|---|---|---|---|---|
| Arts stream students | 0.705 | 0.686 | -0.019 | 45 |
| | (0.181) | (0.178) | (0.054) | |
| Commerce stream students | 0.201 | 0.218 | 0.017 | 45 |
| | (0.105) | (0.125) | (0.034) | |
| Science stream students | 0.094 | 0.096 | 0.002 | 45 |
| | (0.135) | (0.157) | (0.044) | |
| Girls school | 0.782 | 0.773 | -0.009 | 45 |
| | (0.421) | (0.429) | (0.127) | |
| Co-educational school | 0.175 | 0.182 | 0.007 | 45 |
| | (0.387) | (0.395) | (0.117) | |
| Boys school | 0.043 | 0.045 | 0.002 | 45 |
| | (0.209) | (0.213) | (0.063) | |
| N of 12th graders | 180.696 | 168.773 | -11.923 | 45 |
| | (81.100) | (66.324) | (22.143) | |
| N of students per section | 39.496 | 40.404 | 0.908 | 45 |
| | (8.215) | (7.655) | (2.370) | |
| N of sections | 4.957 | 4.818 | -0.138 | 45 |
| | (1.581) | (1.402) | (0.446) | |
| Observations | 23 | 22 | 45 | |

Standard errors are shown in parentheses. Source: Baseline data 2023 and administrative data 2023.

*5.2. Descriptive statistics*

Before analyzing the main causal estimates of impact of the CEP, we present descriptive statistics on what students plan to do and what students actually do after finishing secondary school in both the treatment and control groups, based on data from the baseline, endline, and follow-up surveys.[14] Figure 3 illustrates students' primary and secondary plans after school as stated during the baseline and endline surveys. Figure 4 visualizes the activities students report spending the most and second-most time on during the

[14]Since the figures show only mean values across the two groups without accounting for clustering or stratified randomization, we refrain from commenting on the significance of differences between treatment and control groups.

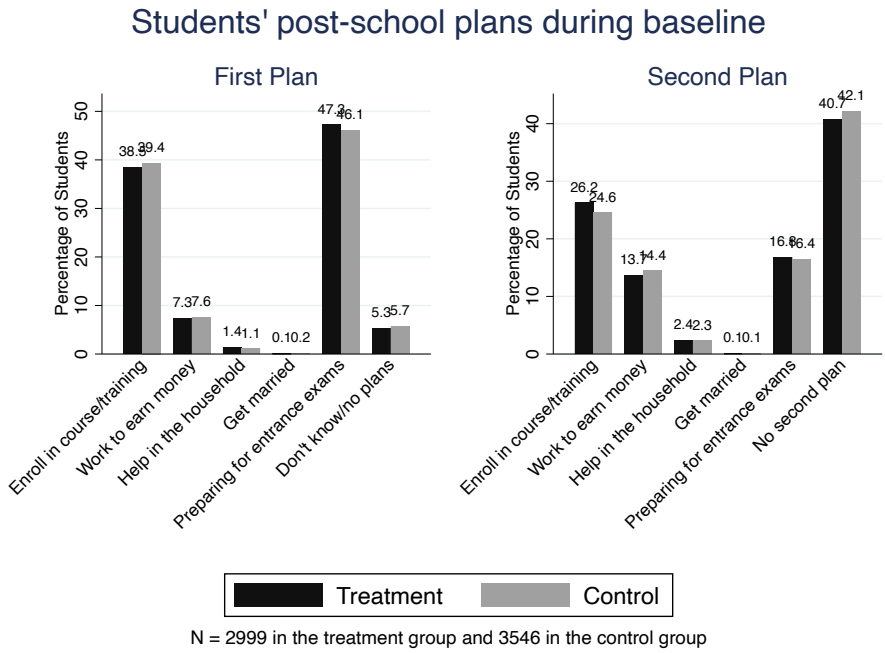
follow-up survey.

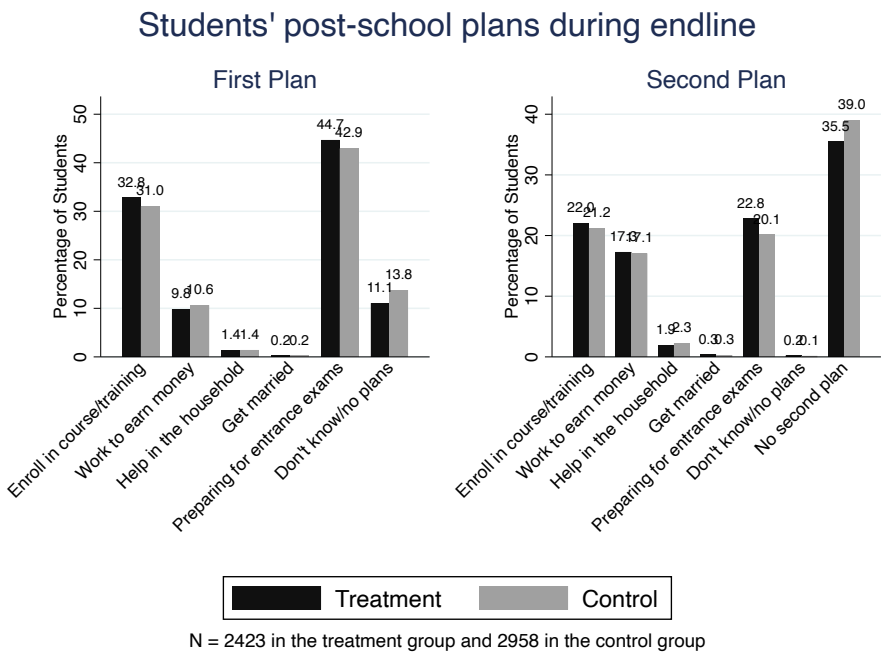
**Baseline – Plans:** At baseline, most students report that their first plan is to prepare for an entrance exam after finishing school (46.1–47.3%), followed by enrolling in a course or training (38.5%-39.4%), and working to earn money (7.3-7.6%). Between 5.3–5.7% of students report not knowing or not having a first plan. The most frequently stated second plans are enrolling in a course or training (24.6–26.2%), preparing for an entrance exam (16.4–16.8%), and working to earn money (13.7–14.4%). 40.7–42.1% of the students report that they do not have a second plan.

**Endline – Plans:** At endline, preparing for an entrance exam continues to be the most frequently stated first plan (42.9–44.7%), followed by enrolling in a course or training (31.0–32.8%) and working to earn money (9.8–10.6%). The share of students who report not knowing or not having a first plan is 11.1–13.8%, slightly higher in the control group. As for second plans, 21.2–22.0% plan to enroll in a course or training, and 20.1–22.8% plan to prepare for an entrance exam. The share of students who plan to work as a second option increases to 17.1–17.3%, while 35.5–39.0% report not having a second plan—again slightly higher in the control group.

**Follow-up – Actions:** In the follow-up survey, the most time-consuming activity reported by students is being enrolled in a course or training (54.5–57.5%), with a slightly higher proportion in the treatment group. This is followed by helping in the household (15.7–19.3%), which students did not mention in their "plans" and which is lower in the treatment group than in the control group. Additionally, 14.3–14.6% report preparing for an entrance exam, and 9.7-10.6% report working for pay.

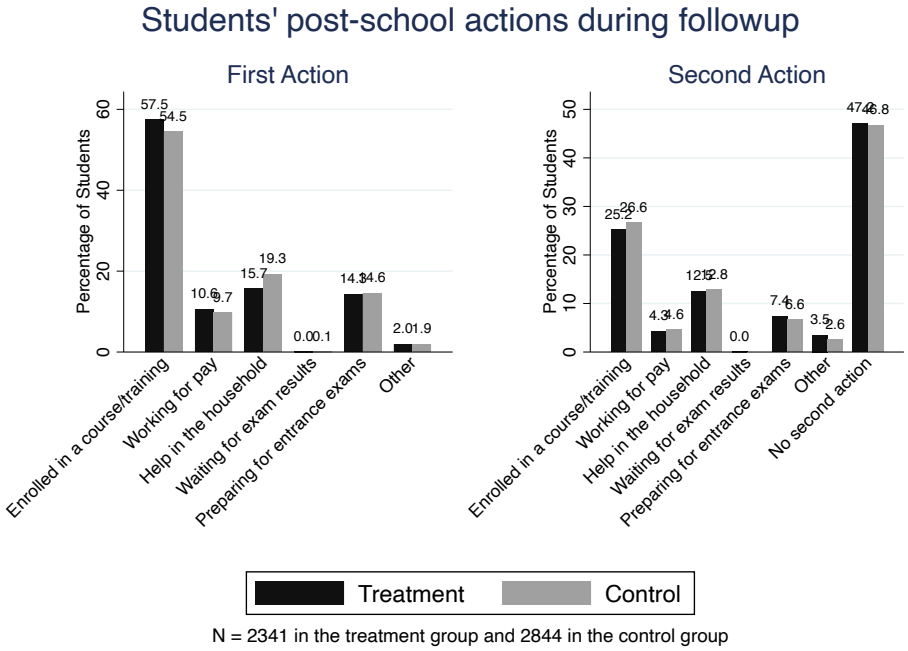
Figure 3: Students' post-school plans as reported in the baseline and endline surveys




Notes: In the baseline and endline, students were asked about their plans after secondary school, and afterwards, whether they had a second plan. Source: Baseline surveys 2023 and endline surveys 2023.

Figure 4: Students' post-school actions as reported in the follow-up surveys




Notes: In the follow-up survey, students were asked about the activity they spend most and second-most of their time on. Source: Follow-up surveys 2024.

When asked about the second-most time-consuming activity, most students (46.8–47.2%) do not report any. Among those who do, the most frequently mentioned activities are being enrolled in a course or training (25.2–26.6%), helping in the household (12.5–12.8%), preparing for entrance exams (6.6–7.4%), and working for pay (4.3-4.6%).

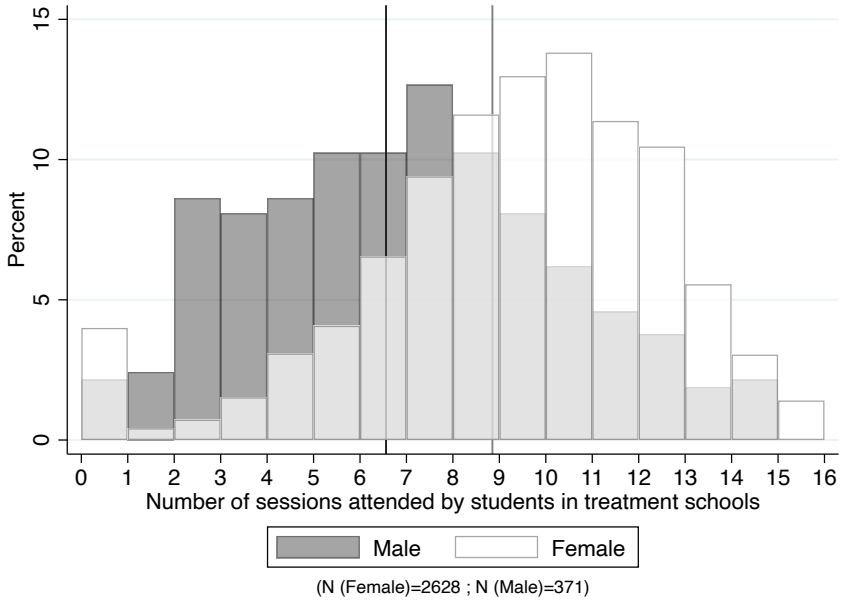
*5.3. Main results*

As pre-registered, we examine our hypotheses on students' future plans, whether they consider interests important for choosing a career path, and their beliefs in intergenerational mobility (short-term) and their actual choices (medium-term). We present our main results on the ITT estimates as pre-registered and additionally from LATE estimates, capturing the impact for students who participated regularly.

Despite the implementation during school hours, students' exposure to the program varied substantially and primarily depended on head teachers giving sufficient time and students' school attendance. On average, in the

treatment group, girls attended 8.8 sessions, while boys attended 6.5 sessions (Figure 5). We define "regular participation" as attending 10 sessions or more, corresponding to the median value across all students.[15]

Figure 5: Session attendance by gender in the treatment group




Notes: This graph visualizes the variation in session attendance among male and female students, with the blue dashed line indicating the mean session attendance of 6.5 sessions for male students and the mean session attendance of 8.8 sessions for female students. Source: Implementation data 2023.

**Short-term outcomes**

Table 4 presents our results for five pre-specified outcomes collected in the endline survey after the completion of the intervention and accounts for baseline covariates as pre-registered.

Panel A of Table 4 reports ITT estimates. Consistent with our hypotheses, we find that students in treatment schools are more likely to plan to take an entrance exam for higher education (H2) and to give importance to their interests when choosing a career (H4). Specifically, students from treated

[15]While we mentioned school attendance as a relevant factor in our study design, we did not pre-register the estimation of the LATE because the facilitators' capacity to collect accurate attendance data was uncertain at the time of pre-registration.

schools are 3.7 percentage points more likely to report planning to sit for an entrance exam ($p < 0.05$), relative to a control mean of 57%. Likewise, they are 7.1 percentage points more likely to report that they consider their interests in occupational choice ($p < 0.01$), relative to a control mean of 44%.

In addition, we find suggestive evidence that the treatment increased students' beliefs in intergenerational mobility (H5). The estimated effect is 3.2 percentage points ($p < 0.10$), compared to a control group mean of 53%. However, we do not detect a statistically significant impact on students' plans to enroll in a course or training program (H1) or their stated likelihood of working for pay in the future (H3). The latter result is unsurprising, given that nearly all students in the control group (96%) see themselves working for pay in the future.

Panel B of Table 4 shows that the estimated effects are larger in magnitude for regular participants. Students with high program exposure are 7.8 percentage points more likely to plan for an entrance exam ($p < 0.05$), 14.8 percentage points more likely to choose an occupation based on their interests ($p < 0.01$), and 6.7 percentage points more likely to believe in intergenerational mobility ($p < 0.05$). As before, we find no statistically significant effects on plans to enroll in training (H1) or to expect working for pay in the future (H3). These results remain qualitatively unchanged after adjusting for multiple hypothesis testing, as shown in Figure F1 in the Appendix. The only exception is the effect on belief in intergenerational mobility, which becomes suggestively significant ($q = 0.086$). Taken together, these results provide evidence that the CEP successfully shifted students' aspirations and beliefs in the short run, particularly among students who were regularly exposed to the content.[16]

[16] Although survey outcomes are self-reported, demand effects cannot be completely ruled out, particularly as students may have internalized the program's emphasis on planning and interest-based choices.

Table 4: Short-term results - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) |
|---|---|---|---|---|---|
| ***Panel A: ITT*** | | | | | |
| CEP | 0.020 | 0.037** | 0.001 | 0.071*** | 0.032* |
| | (0.015) | (0.016) | (0.006) | (0.016) | (0.016) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 |
| ***Panel B: LATE*** | | | | | |
| Regular CEP student | 0.041 | 0.078** | 0.003 | 0.148*** | 0.067** |
| | (0.030) | (0.034) | (0.012) | (0.033) | (0.033) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.51 | 0.57 | 0.96 | 0.44 | 0.53 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

**Medium-term outcomes**

Next, we examine whether the observed short-run impacts on students' plans and beliefs translate into measurable behavioral changes in the medium term. Approximately 10-12 months after the program implementation, we re-contacted the students to collect information on actual post-secondary-school behaviors, labor market participation, and some non-cognitive outcomes. Table 5 presents the results for four pre-registered and one exploratory outcome. In addition to having enrolled in a course, having taken an entrance exam, having completed a course and working for pay, we include the exploratory outcome of whether students report spending time in household work which in Figure 4 appeared as an activity a relevant share of students is spending time on with 15.7-19.3% of the students describing it as their most time-consuming activity and 12.5-12.8% as their second-most time-consuming activity despite of not planning for it.

Similar to the short-term results, panel A of Table 5 shows the ITT estimates. Across all prespecified outcomes, including enrollment in a course or training (H6), planning to take or having taken an entrance exam (H7), course completion (H8), and currently working for pay (H9), we do not find

any statistically significant effects. All point estimates are small in magnitude and not distinguishable from zero at the conventional significance levels. The only significant difference between treatment and control group students appears in spending time on household work, with students from treated schools being 3.1 percentage points ($p < 0.05$) less likely to do so compared to control group students, which corresponds to a reduction of 10%, considering the control group mean of 32%.

Panel B shows that this pattern also holds in the LATE estimates. For students with high exposure to the CEP, we again find no evidence of positive treatment effects on any of the pre-registered medium-term outcomes. The estimates remain close to zero and statistically insignificant across all specifications. The negative impact on spending time on household work increases to 6.8 percentage points ($p < 0.05$), corresponding to a reduction of 21%, considering the control group mean of 32%.

Table 5: Medium term results - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|
| ***Panel A: ITT*** | | | | | |
| CEP | -0.008 | 0.007 | 0.007 | 0.003 | -0.031** |
| | (0.014) | (0.020) | (0.016) | (0.009) | (0.013) |
| N | 5244 | 5244 | 1652 | 5244 | 5244 |
| ***Panel B: LATE*** | | | | | |
| Regular CEP student | -0.018 | 0.015 | 0.017 | 0.008 | -0.068** |
| | (0.029) | (0.044) | (0.039) | (0.018) | (0.028) |
| N | 5244 | 5244 | 1652 | 5244 | 5244 |
| Control mean | 0.70 | 0.29 | 0.21 | 0.10 | 0.32 |

*Notes:* As pre-registered, except the last outcome in column 5. All specifications include baseline covariates. Only students who were not enrolled in any course or training were asked if they had completed a short type of course or training in the last three months after 12th grade exams. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Finally, we turn to two pre-registered non-cognitive outcomes that we hypothesized to be impacted by the CEP — locus of control and self-efficacy. We measured both in the endline and in the follow-up survey, and Table 6 shows that the CEP does not seem to impact these non-cognitive outcomes, with all estimated to be close to zero and insignificant. The absence of effects on locus of control and self-efficacy may suggest that these internal psychological traits are less malleable in the short term or require more intensive or sustained interventions.

Accounting for multiple-hypothesis testing also in the medium-term does largely not change our results, except for the impact of the CEP reducing the likelihood of students helping in the household now being marginally insignificant ($q = 0.112$ for ITT and $q = 0.107$ for LATE) as shown in Figure F1 in the Appendix.

Table 6: Short-term and medium-term results on secondary hypotheses - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Locus of control Short-term (H10) | (2) Self efficacy Short-term (H11) | (3) Locus of control Medium-term (H12) | (4) Self efficacy Medium-term (H13) |
|---|---|---|---|---|
| ***Panel A: ITT*** | | | | |
| CEP | 0.016 | 0.003 | 0.012 | -0.002 |
| | (0.016) | (0.008) | (0.014) | (0.005) |
| N | 5381 | 5381 | 5187 | 5187 |
| ***Panel B: LATE*** | | | | |
| Regular CEP student | 0.033 | 0.007 | 0.025 | -0.004 |
| | (0.032) | (0.015) | (0.031) | (0.011) |
| N | 5381 | 5381 | 5187 | 5187 |
| Control mean | 0.41 | 0.92 | 0.38 | 0.93 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

In summary, we find that the CEP significantly impacted several dimensions of students' aspirations and beliefs in the short term, particularly strongly for students with relatively high exposure to the program. These

include planning to take entrance exams, prioritizing personal interests in occupational choices, and holding stronger beliefs in intergenerational mobility. Yet, these attitudinal and aspirational shifts did not translate into observable behavioral outcomes overall consistent with a growing literature that documents a gap between aspirational change and actual behavior, especially when structural barriers remain unaddressed and parental support is insufficient (Genicot and Ray, 2020, 2017; Carrell and Sacerdote, 2017; Kipchumba et al., 2024).

One potential explanation for the lack of impact on students' behavior is that while the intervention succeeded in 'unlocking students' minds' in terms of their plans, it did not unlock the binding constraints they continue to face that prevent them from following through in terms of behavioral change. Examples include financial barriers, limited mobility, and gender norms, which may have limited their decision-making power and their ability to enroll in tertiary education and skill formation.[17] Another potential explanation is that the follow-up survey was conducted too early after leaving school and could therefore not fully realize educational or labor market decisions, particularly for those involving multiple steps, including entrance exams and waiting times. Finally, the design of the CEP may not have included sufficient components to support long-term follow-through (e.g., financial aid navigation, parental engagement, or mentorship), which could have been critical for interventions targeting the transition from secondary to tertiary education.

### *5.4. Mechanisms*

To assess the relevance of the hypothesized mechanisms, Table 7 presents regressions of each pre-registered mediator on the treatment indicator and baseline covariates, following equation 6. Although we defined the coding

[17] According to the 2022–23 Household Consumption Expenditure Survey, average monthly per capita expenditure in urban India is INR 6,459 (INR 77,500 annually) (Ministry of Statistics and Programme Implementation (MOSPI), 2024). Although the government-run Industrial Training Institutes have a relatively low fee (INR 1,600 - 1,900 per year), this excludes indirect costs such as transport, safety concerns, or student accommodation. Moreover, government institutes often face constraints in terms of limited places, lower perceived reputation, and accessibility, making private providers more likely to be chosen. For households in the lower-middle part of the distribution, such as our sample, a typical vocational or training program costs around INR 10,000 - 25,000 per year, corresponding to several months of per capita consumption.

of the mechanisms in the preregistration, we adjusted the coding of these mechanisms, but followed the same logic and used the same mechanisms. All mechanisms are coded the same way as the preregistration, except for information, educational aspirations, and salary aspirations as detailed in Appendix C. The results indicate that the program effectively alleviates information constraints and encourages students to view facilitators as sources of information and inspiration. However, it does not raise education or salary aspirations, increase discussions of career plans with peers, or encourage students to follow peers' actions. Thus, two of the hypothesized mediators are positively impacted by the treatment, and these are also the ones that explain most of the treatment effects on the outcomes, as shown in Figure 6 when we turn to the decomposition of the short-term impacts.

Table 7: Mediators - ITT estimates accounting for students' attendance - with baseline covariates

| | (1) Information | (2) Educational asp. | (3) Salary asp. | (4) Peers comm. | (5) Peers influence | (6) Community support | (7) Facilitators |
|---|---|---|---|---|---|---|---|
| CEP | 0.124*** | -0.008 | 0.019 | -0.008 | -0.092 | 0.018 | 0.132*** |
| | (0.021) | (0.015) | (0.019) | (0.014) | (0.116) | (0.011) | (0.012) |
| N | 5381 | 5352 | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.49 | 0.74 | 0.49 | 0.76 | 2.87 | 0.84 | 0.05 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

To understand how the CEP causes the observed impacts, we decompose the total treatment effect using a causal mediation framework. Following Imai et al. (2010) and Heckman and Pinto (2015), we estimate the indirect effects through each mediator and the remaining direct effect.

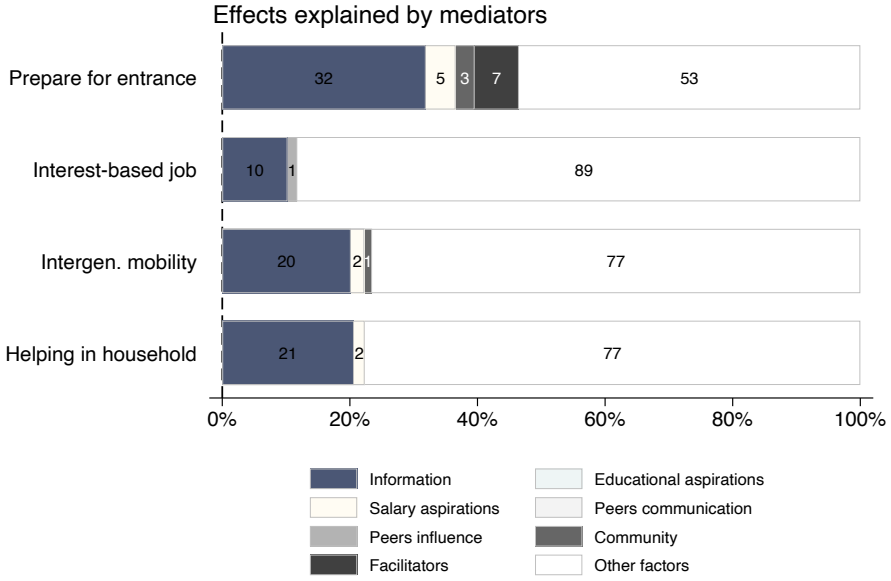
Figure 6 summarizes the share of the treatment effect explained by each mediator for three key outcomes positively affected by the CEP: (i) planning to take an entrance exam, (ii) interest in pursuing a job aligned with one's preferences, (iii) beliefs about intergenerational mobility, and (iv) spending time in doing household work. The decomposition relies on average causal mediation effects estimated separately for each mediator, using post-treatment mediator values. The stacked bars display the contribution of each

mediator alongside the residual "unexplained" portion of the effect.

The hypothesized mediators explain 47% of the treatment effect on preparing for entrance exams, but only 11% and 23% of the impact on interest-based job selection and beliefs about intergenerational mobility, respectively. In terms of medium-term effects on helping in the household, although not preregistered, the mediators explain 23% of the impact. As expected, alleviating information constraints is the dominant mechanism, followed by facilitator effects.

For preparing for entrance exams, 32% of the effect is explained by reduced information frictions, 5% by higher salary aspirations, 3% by community support, and 7% by viewing facilitators as inspirational sources. For interest-based job selection, 10% of the effect is explained by information frictions, and 1% by peers' influence. Regarding intergenerational mobility beliefs, 20% is explained by alleviating information constraints, 2% is explained by higher salary aspirations, and 1% by community support. Lastly, for helping in the household, 21% is explained by reduced information frictions and 2% by higher salary expectations.

Figure 6: Mechanisms from mediation analysis




Notes: This graph visualizes the relevance of the hypothesized mechanisms for explaining the significant short-term impacts of the CEP. Source: Baseline survey 2023 and endline survey 2023.

Overall, these results suggest that reducing information frictions and providing relatable facilitators are the two most important channels through which the program operates. Nevertheless, raising salary aspirations and community support also contribute to explaining the impacts. In contrast, peer discussions appear to play no role in explaining the observed short-term or medium-term impacts.

### *5.5. Heterogeneous treatment effects*

In our analysis of heterogeneous treatment effects, we examine whether impacts differ depending on the gender norms students face at home, their mobility, and their beliefs in meritocracy. As outlined in Section 4.4, we hypothesized that female students from non-patriarchal households, students with greater mobility, and students with stronger meritocratic beliefs would benefit more from the CEP. Figure 7 plots the coefficients from ITT regressions for short-term (left panel) and medium-term (right panel) outcomes with the corresponding regressions in Appendix G.

*Patriarchal norms at home*

We examine whether the effects of the CEP differ for female students based on the gender norms in their household, proxied by whether any woman has worked outside the home in the last two years.[18] The top panel of Figure 7 plots the estimates for short and medium-term outcomes with the corresponding regressions in Tables G1 and G2, respectively. While we cannot reject the null hypothesis of equal effects for the two subgroups when testing the estimated treatment effects against each other, we observe larger and statistically significant effects among female students from households with more patriarchal norms, particularly on planning to enroll in a course or training, planning to take an entrance exam, and believing in intergenerational mobility. The effects for female students from non-patriarchal households are insignificant, except for the CEP increasing the attention given to interests in the job selection for both groups.

[18] The presence of women working for pay does not necessarily indicate progressive gender attitudes, as it may also reflect financial necessity. However, considering two households living in similar economic conditions as captured through the asset index and number of household members, a household with a woman working outside the home can reasonably be considered less patriarchal. We account in all specifications for the household's asset index and household size as reported by the student.

This pattern continues for the medium-term outcomes, with female students from non-patriarchal households being more likely to have completed a short course due to the CEP and more likely to have reduced the time they spend helping in the household. In terms of suggestive evidence, the CEP slightly increases the likelihood of girls from non-patriarchal households working for pay. Yet, as before, we cannot reject the null-hypothesis on equal effects for the two subgroups.

One plausible explanation for this result is that students from less patriarchal households already have women around them engaged in the labor force and receive more advice and support at home. This may enable them to continue with skill formation independent of their exposure to the CEP. This can also be seen in Tables G1 and G2, where students from non-patriarchal households are more likely to enroll in a course or have completed a course irrespective of the treatment. However, this reasoning does not hold for the differences between girls from patriarchal and non-patriarchal households with respect to planning to take an entrance exam.

*Mobility*

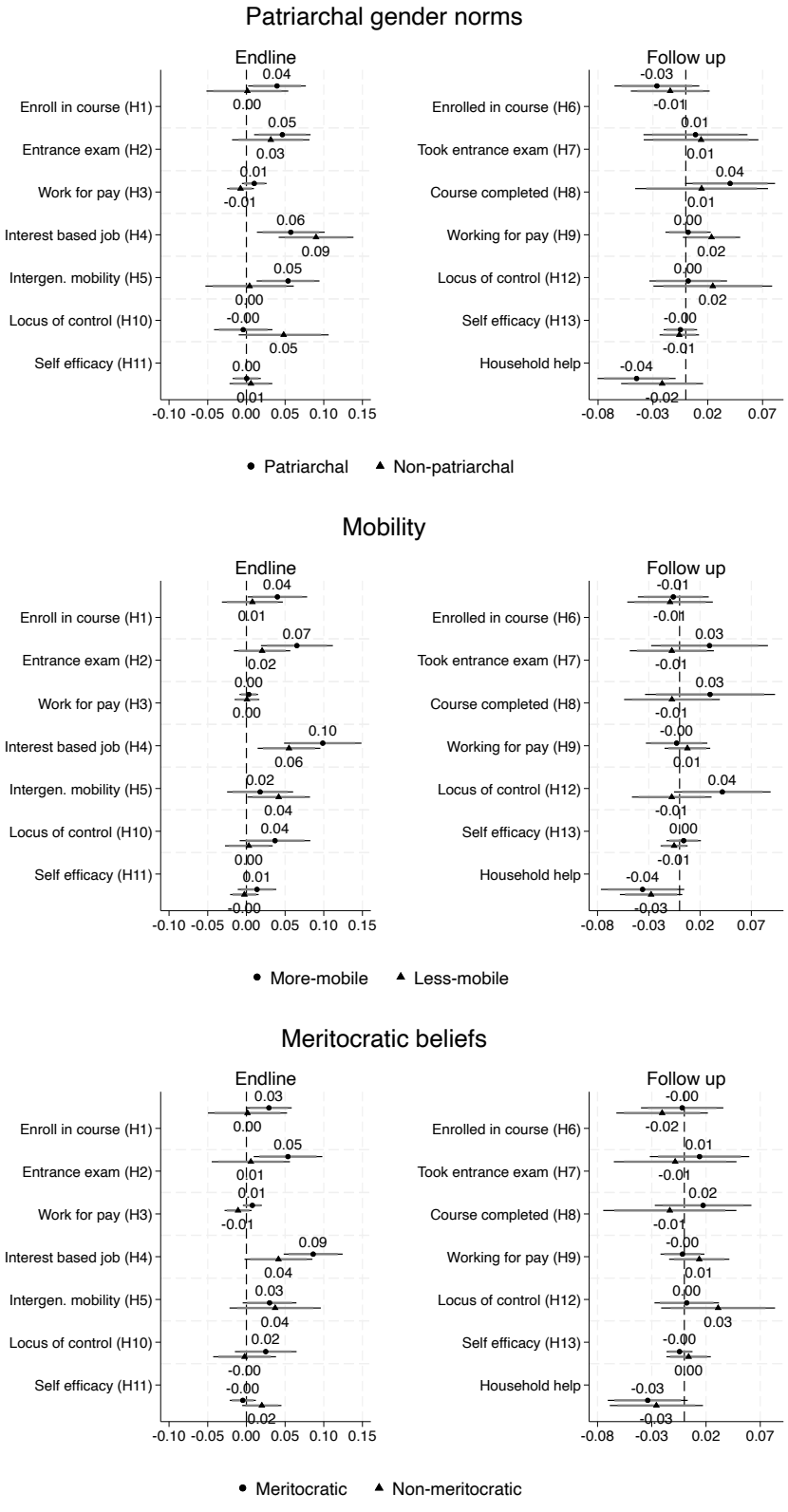
We examine whether the impacts of the CEP intervention vary by students' baseline mobility, shown in the middle panel of Figure 7. The corresponding regressions in the appendix in Tables G3 and G4 show the heterogeneity estimates for short- and medium-term outcomes, respectively, comparing more-mobile and less-mobile students. As expected, the CEP significantly impacts more mobile students with respect to their educational plans. Due to the CEP, they are more likely to plan to enroll in a course and to plan to take an entrance exam than students in the control group. The impact on planning to take an entrance exam is significantly weaker for less mobile students. The CEP positively impacts interest-based job selection for more mobile and less mobile students, with a more pronounced estimate for more mobile students. In contrast, the impact of the CEP on intergenerational mobility is driven by less mobile students who generally have weaker beliefs in intergenerational mobility, as shown in G3, irrespective of the treatment.

In the medium term, differences in treatment effects by mobility status are smaller and largely statistically insignificant. We continue to see marginally significant impacts on locus of control for more-mobile students, suggesting some sustained non-cognitive gains for this subgroup, but not significantly different from the impact on less-mobile students. The reduced time spent helping in the households is marginally significant for both groups.

*Meritocracy*

*Meritocracy*: We examine whether the impacts of the CEP vary by students' beliefs in meritocracy, shown in the last panel of Figure 7. Tables G5 and G6 in the appendix show the corresponding regressions for short- and medium-term outcomes, respectively. As expected, we do observe more pronounced impacts in the short term for meritocratic students. CEP significantly increases the likelihood of meritocratic students planning to enroll in a course, planning to take an entrance exam, and aiming for an interest-based job. The positive impact on interest-based job selection is significant for both groups, but seems to be larger for the meritocratic students than for the non-meritocratic students. Apart from this, the treatment effects for the non-meritocratic students are mostly small and insignificant.

Figure 7: Heterogeneity - marginal effects for subgroups I



There is little evidence of differential impacts by meritocratic beliefs in the medium term, with one exception. The average impact of reducing time spent on household work is significant for the meritocratic students and insignificant for the non-meritocratic ones, but the difference between the two estimates is insignificant.

Additional heterogeneity analyses reveal that the CEP has different impacts across student subgroups (see Appendix H for estimations).

**School type:** The CEP increases the share of students planning to take an entrance exam only among those attending single-sex schools, with no such effect observed among students in co-educational settings. However, in the medium term, the program leads to an increase in course enrollment only among students from co-educational schools, not those from single-sex schools.

**Stream:** When examining differences by stream, the CEP significantly increases the likelihood of interest-based career selection among Arts students but not among those from non-Arts streams. However, no difference in treatment effect between the two groups was observed.

**Religion/caste:** Religious affiliation or caste belonging also moderates the impact of the CEP. In the medium term, the program raises the proportion of non-Hindu girls working for pay, with no comparable effect observed for Hindu girls. This finding may reflect enhanced economic empowerment, though it might also suggest a premature transition into the labor market that limits potentially more beneficial long-term skill development. Further, we also find that the CEP significantly increases the likelihood of interest-based career selection in the short term among upper caste students but not among non-upper caste students.

**Grades:** Academic performance further differentiates program impacts. The CEP increases the likelihood of planning to take an entrance exam among higher-performing students, but not among lower-performing peers. While the program encourages students across the performance spectrum to pay greater attention to their interests, the effect is significantly stronger among lower-performing students. Finally, the program's positive impact on beliefs about intergenerational mobility is significant only for higher-performing students.

**Gender:** Finally, while not pre-registered due to the relatively small sample of male students (about 10%), we also examine heterogeneity by gender expecting more positive impacts for female students as they tend to be more disadvantaged. Appendix Table H6 shows that the CEP increases

the probability that female students report the completion of short courses and working for pay in the follow-up survey significantly more than for male students suggesting that for female students positive changes in intentions translated into some behavioral change as well.[19]

*5.6. Additional impacts - exploratory*

In addition to the pre-registered outcomes, the data allows us to examine other outcomes that closely relate to students' career readiness. Table 8 shows that the CEP increases the share of students thinking about their future, knowing their desired occupation, feeling confident about achieving their career objectives, and telling parents about their career objectives significantly, as reported in the endline surveys. Particularly strong are the impacts on students knowing their desired occupation and feeling confident about it in the short term. In addition, we also observe that students from treatment school continue to state that they consider their interests an important factor when choosing a career path. Accounting for multiple hypothesis testing, these results remain significant ($p < 0.05$) except for interest-based job selection, which becomes marginally insignificant ($p = 0.112$ for ITT and $p = 0.107$ for LATE) as shown in Figure F1 in the Appendix.

[19] A potential downside of the positive impact on female students is that CEP appears to reduce the share of male students reporting completion of a short course or training. However, at least suggestively, their likelihood of being enrolled in a course at the time of the follow-up survey is higher. Given the relatively small sample size of male students, we interpret these results with caution.

Table 8: Additional outcomes - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Thought about future (E1) | (2) Knows occu-pation (E2) | (3) Confident about objective (E3) | (4) Tells parents (E4) | (5) Interest-based job (Follow-up) (E5) |
|---|---|---|---|---|---|
| ***Panel A: ITT*** | | | | | |
| CEP | 0.029*** | 0.062*** | 0.061*** | 0.030** | 0.034** |
| | (0.010) | (0.013) | (0.012) | (0.013) | (0.015) |
| N | 5381 | 5334 | 5334 | 5334 | 5203 |
| ***Panel B: LATE*** | | | | | |
| Regular CEP student | 0.060*** | 0.129*** | 0.127*** | 0.063** | 0.075** |
| | (0.023) | (0.028) | (0.027) | (0.027) | (0.034) |
| N | 5381 | 5334 | 5334 | 5334 | 5203 |
| Control mean | 0.89 | 0.74 | 0.73 | 0.65 | 0.63 |

*Notes:* Exploratory analysis. All specifications include baseline covariates. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

## 6. Conclusion

This study contributes important evidence on the role and limitations of low-cost, scalable career guidance programs in supporting young women to transition from secondary education to further skill formation in urban India. The Career Exploration Program (CEP) evaluated here places students, especially young women from disadvantaged backgrounds, at the center of their career decision-making processes. Unlike conventional interventions that primarily offer one-way information or advice, counselors give students the CEP, which emphasizes self-reflection, peer discussions, and exposure to relatable facilitators to help students understand their interests, explore options, and develop actionable career plans.

We find that the CEP significantly improves students' short-term aspirations and addresses the information constraints that students face. Students in the treatment group were more likely to report plans to take entrance exams, to prioritize personal interests in career decisions, and to believe in intergenerational mobility. Although not pre-registered, we also find that the CEP

improves students' knowledge and confidence about what they would like to do in the future. These effects were stronger among students who attended at least ten sessions, highlighting the importance of regular program exposure and the critical role of schools in supporting implementation. Our mechanism analysis suggests that these improvements were primarily driven by increased access to career-related information and students viewing facilitators as key sources of guidance and inspiration. These findings demonstrate that well-designed career guidance programs, even in resource-constrained environments, can raise students' aspirations and promote greater self-awareness regarding their educational and professional trajectories.

However, we find no average impacts on longer-term educational enrollment or training behavior within a one-year post-intervention follow-up period. While these attitudinal changes observed in the short term did not translate into behavioral changes on average, our heterogeneity analysis reveals that the CEP is particularly effective in raising educational aspirations for some subgroups, such as girls from more patriarchal households, students who are more mobile, and students who hold more meritocratic beliefs, with few treatment impacts continuing in the medium term. As such, co-educational students are more likely than single-sex students to benefit from the CEP in terms of increasing their enrollment rates in longer-term courses, and non-Hindu students are more likely to engage in work for pay due to the CEP than the Hindu students in the treatment group.

Taken together, these findings underscore a key challenge in career guidance interventions. While changing mindsets and intentions is necessary and such interventions have been successful in different contexts, it is not always sufficient to produce behavioral change in the longer term, particularly in contexts where structural and normative barriers persist and constrain female and male students in different ways.

The absence of longer-term impacts may be attributable to several factors. First, systemic barriers—including financial constraints, rigid gender norms, and mobility restrictions can inhibit students' ability to act on their aspirations. Second, the relatively short follow-up period about six months after the final secondary school exams may not fully capture delayed educational decisions or multi-step transitions. While treatment and control group students are equally likely to be enrolled, there could be future differences in completion rates, study performance, or satisfaction with a chosen career path, which are not captured here. Third, the CEP may have lacked complementary supports such as more guidance on financial aid, con-

tinued mentorship, or parental engagement, which are potentially needed to sustain behavioral change beyond school. Fourth, students may face issues with newly introduced procedures or regulations for access to higher education in India, such as the introduction of centralized exams for university admission, which can be administratively complicated for the students and potentially require more preparation. Lastly, it could also be the presence of demand effects: students exposed to the CEP may have reported more forward-looking plans because they internalized the program's emphasis on planning and interest-based choices, rather than because of actual behavioral change. 1However, the stronger LATE estimates suggest that students who had more exposure to the CEP have stronger effects on their plans, suggesting that the program's content mattered and not just the exposure to the program.

These findings carry several implications for both policymaking and future research. From a policy perspective, scalable programs like this locally designed CEP can be a valuable component of broader strategies to address low female labor force participation. Yet, to translate aspirations into action, such interventions should be integrated into efforts that also address material and social constraints. For example, pairing career guidance with financial planning workshops, parental involvement, or digital mentorship platforms may increase their long-term effectiveness. Future research should extend the follow-up horizon to capture better delayed impacts, particularly when career pathways are fragmented and decision-making unfolds over time. Understanding how the characteristics of facilitators and the degree of parental involvement influence outcomes could also offer insights for optimizing program design.

In conclusion, while this study shows that low-cost, school-based career guidance programs can lead to meaningful short-term shifts in planning and aspiration, translating those gains into sustained behavioral change will likely require complementary supports that address the broader set of constraints that young women in India and similar contexts face. As policymakers consider expanding such programs, careful attention to these contextual realities will be critical to unlocking their full potential.

**Declaration of generative AI and AI-assisted technologies in the writing process**

During the preparation of this work, the authors used ChatGPT to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Appendix

### Appendix A. Timeline

2022: Piloting of surveys, randomization, and implementation
May-July 2023: Baseline surveys
June-November 2023: Implementation of the CEP in schools
November-December 2023: Endline data collection in schools
April-May 2024: Final senior secondary school exams
August 2024: Enrollment for higher education courses takes place
October-November 2024: Follow-up 1 data collection

### Appendix B. Intervention description

The CEP envisions empowering young adults (15-21years) to pursue interest-aligned learning and career opportunities. It aims to equip students with the skills, knowledge, mindsets, and a clear plan that enables them to feel confident about working towards a career they choose. It enables students to explore career options aligned with their interests, experience them through tasks, and prepare a plan of action toward reaching their career goals. The program takes a student through a 5-step process leading them to make a choice aligned with their interests and skills.

1. **Start to think!** - Thinking early about making a choice lets one take charge of their career pathway, giving time at hand to explore.

2. **Think openly!** - Not limiting yourself to a few conventional careers, but thinking openly, e.g., thinking of the healthcare sector instead of just being a doctor, opens one up to a variety of options to choose from.

3. **Understand yourself!** - Experiencing different types of work, getting feedback, and reflecting on those experiences to find which careers align more with yourself.

4. **Prepare to succeed!** - Creating a plan of action to reach the chosen goal, including the research of where to go and how, and steps to prepare oneself for the same.

5. **Thrive in your chosen path!** - Taking continuous action- reflection cycles to up-skill and succeed.

Throughout these five stages, the program uses a comprehensive virtual portal designed to provide students with a wholesome learning experience about careers. It hosts career information in the form of courses.

1. **Career expert video** – A short video of a contextual role model sharing their experience about the work done in the career, skills required on the job, and how to prepare for it

2. **Career path** - Broad pathway consisting of courses, internships, and jobs to reach an average salary of Rs. 30k-40k

3. **Quizzes** - Every learning content is followed by a small quiz (4-5 objective questions) to check students' understanding and create an opportunity for the student to recollect their learning

4. **Home challenge** - A home-based activity designed to give students an experience of what work in a career would look like. It is based on the core skills required for the job.

5. **Reflection form** - Set of detailed reflective questions to help students synthesize their experience of the career explored.

Apart from the 50+ career courses currently hosted on the portal, courses on introducing the concept of career exploration, creating a plan of action, and entrepreneurial skills like communicating your idea, planning your finances, navigating uncertainties, etc., are also available to help students build the necessary mindset.

These videos have been taken from interviews of high-achieving personalities like Sudha Murthy, Ruma Devi, and Rohit Thakral, who are initially from financially constrained backgrounds similar to those of the students participating in the program. The virtual portal allows students to pursue independent exploration at their own pace after registration.

## Appendix C. Deviation from pre-registration

*LATE estimation*

Our pre-analysis plan mentioned that incomplete school attendance is a relevant problem in our setting. Observing the variation in the number of sessions students participated in motivates us to examine the LATE in addition to the ITT. We did not plan for this at the time of the pre-registration because attendance data had to be collected by the implementing organization, and we could not be sure about the facilitators' capacity to collect accurate attendance data.

*Mediators*

In the pre-registration, the information variable was defined as a binary indicator equal to 1 if a student had complete knowledge of at least one of the three careers they wished to pursue, and 0 otherwise. However, in the current analysis, we adapt this definition by standardizing the variable relative to the median level of information within the sample. Under this revised coding, students who report more knowledge about a broader range of careers than the median respondent are assigned a value of 1, while those with less information are coded as 0. A similar approach is applied to the educational and salary aspiration outcomes. We use a median split to code students as having "higher" aspirations (coded as 1) if their reported educational or salary goals exceed the sample median, and 0 otherwise. From our point of view, this adjustment is methodologically more appropriate since it captures better the variation across students in terms of information received, educational aspiration, and salary aspiration.

*Heterogeneity*

We planned to examine the following hypothesis "Students who update their beliefs from baseline to endline on gender roles are expected to benefit more from the program.". While we collected all the data for this, this hypothesis suffers from endogeneity since students in treatment schools might be more likely to update their beliefs on gender norms. Therefore, we decided not to include this heterogeneity analysis in our results.

## Appendix D. Results only for girls

Table D1: Short-term results only for girls - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) |
|---|---|---|---|---|---|
| ***Panel A: ITT*** | | | | | |
| CEP | 0.025 | 0.040** | 0.003 | 0.070*** | 0.034* |
| | (0.016) | (0.016) | (0.006) | (0.018) | (0.018) |
| N | 4794 | 4794 | 4794 | 4794 | 4794 |
| ***Panel B: LATE*** | | | | | |
| Regular CEP student | 0.053 | 0.087** | 0.007 | 0.152*** | 0.074* |
| | (0.034) | (0.035) | (0.013) | (0.039) | (0.038) |
| N | 4794 | 4794 | 4794 | 4794 | 4794 |
| Control mean | 0.52 | 0.58 | 0.96 | 0.45 | 0.53 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D2: Medium-term results only for girls - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|
| ***Panel A: ITT*** | | | | | |
| CEP | -0.022 | 0.011 | 0.029* | 0.011 | -0.036** |
| | (0.014) | (0.022) | (0.016) | (0.008) | (0.015) |
| N | 4658 | 4658 | 1459 | 4658 | 4658 |
| ***Panel B: LATE*** | | | | | |
| Regular CEP student | -0.049 | 0.024 | 0.076* | 0.025 | -0.081*** |
| | (0.031) | (0.049) | (0.040) | (0.019) | (0.031) |
| N | 4658 | 4658 | 1459 | 4658 | 4658 |
| Control mean | 0.70 | 0.28 | 0.20 | 0.09 | 0.35 |

*Notes:* As pre-registered, except the last outcome in column 5. All specifications include baseline covariates. Only students who were not enrolled in any course or training were asked if they had completed a short type of course or training in the last three months after 12th grade exams. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D3: Short-term and medium-term results on secondary hypotheses only for girls - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Locus of control Short-term (H10) | (2) Self efficacy Short-term (H11) | (3) Locus of control Medium-term (H12) | (4) Self efficacy Medium-term (H13) |
|---|---|---|---|---|
| ***Panel A: ITT*** | | | | |
| CEP | 0.016 | 0.002 | 0.011 | -0.005 |
| | (0.017) | (0.008) | (0.015) | (0.005) |
| N | 4794 | 4794 | 4605 | 4605 |
| ***Panel B: LATE*** | | | | |
| Regular CEP student | 0.035 | 0.005 | 0.025 | -0.012 |
| | (0.036) | (0.017) | (0.033) | (0.012) |
| N | 4794 | 4794 | 4605 | 4605 |
| Control mean | 0.41 | 0.92 | 0.38 | 0.92 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## Appendix E. Results accounting for attention checks in baseline and endline

Table E1: Short-term results accounting for attention check in baseline and endline - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) |
|---|---|---|---|---|---|
| ***Panel A: ITT*** | | | | | |
| CEP | 0.020 | 0.038** | 0.002 | 0.071*** | 0.032* |
| | (0.015) | (0.016) | (0.006) | (0.016) | (0.016) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 |
| ***Panel B: LATE*** | | | | | |
| Regular CEP student | 0.042 | 0.078** | 0.003 | 0.148*** | 0.066** |
| | (0.030) | (0.034) | (0.012) | (0.033) | (0.033) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.51 | 0.57 | 0.96 | 0.44 | 0.53 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table E2: Medium-term results accounting for attention check in baseline - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|
| ***Panel A: ITT*** | | | | | |
| CEP | -0.008 | 0.007 | 0.007 | 0.004 | -0.031** |
| | (0.014) | (0.020) | (0.016) | (0.009) | (0.013) |
| N | 5244 | 5244 | 1652 | 5244 | 5244 |
| ***Panel B: LATE*** | | | | | |
| Regular CEP student | -0.018 | 0.014 | 0.017 | 0.008 | -0.067** |
| | (0.029) | (0.043) | (0.039) | (0.018) | (0.028) |
| N | 5244 | 5244 | 1652 | 5244 | 5244 |
| Control mean | 0.70 | 0.29 | 0.21 | 0.10 | 0.32 |

*Notes:* As pre-registered, except the last outcome in column 5. All specifications include baseline covariates. Only students who were not enrolled in any course or training were asked if they had completed a short type of course or training in the last three months after 12th grade exams. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

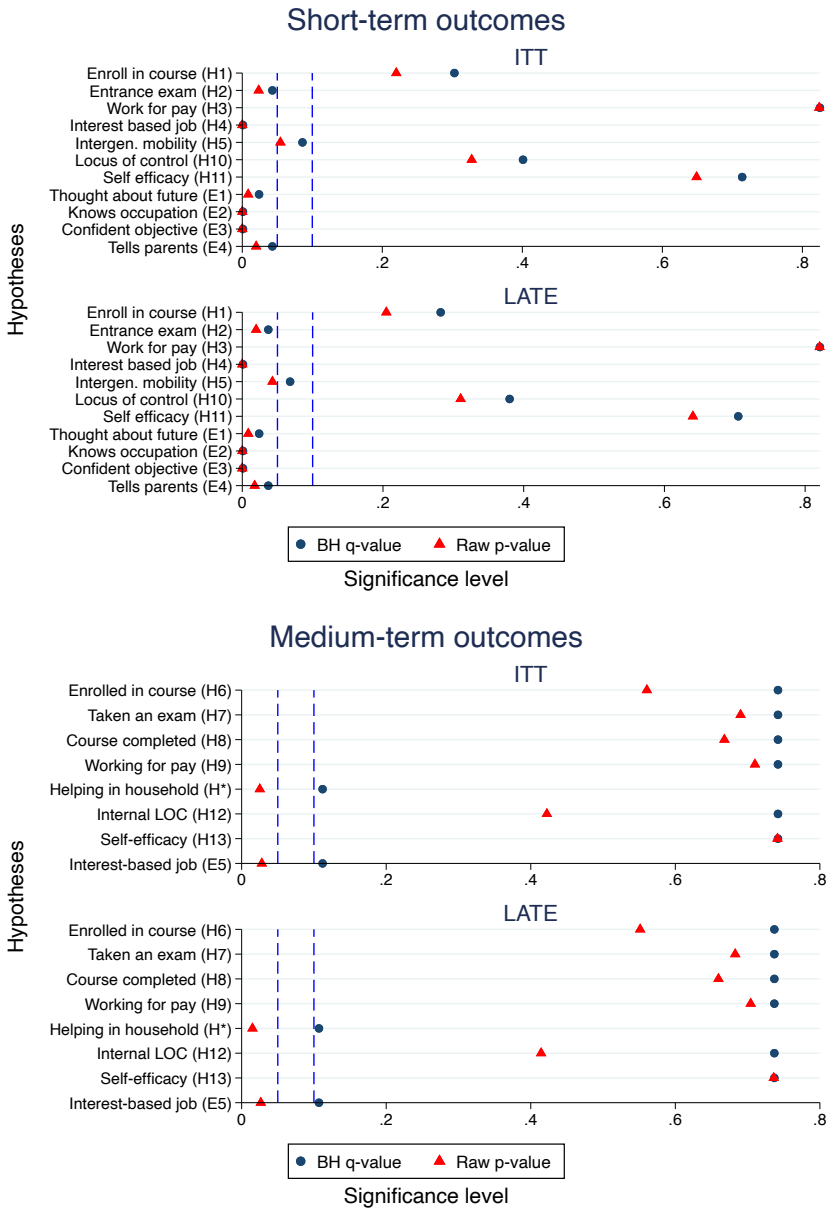
Table E3: Short-term and medium-term results on secondary hypotheses accounting for attention check in baseline and endline for short-term and only baseline for medium-term results - ITT and LATE accounting for students' attendance - with baseline covariates

| | (1) Locus of control Short-term (H10) | (2) Self efficacy Short-term (H11) | (3) Locus of control Medium-term (H12) | (4) Self efficacy Medium-term (H13) |
|---|---|---|---|---|
| ***Panel A: ITT*** | | | | |
| CEP | 0.015 | 0.004 | 0.011 | -0.002 |
| | (0.016) | (0.008) | (0.014) | (0.005) |
| N | 5381 | 5381 | 5187 | 5187 |
| ***Panel B: LATE*** | | | | |
| Regular CEP student | 0.032 | 0.007 | 0.025 | -0.004 |
| | (0.032) | (0.015) | (0.031) | (0.011) |
| N | 5381 | 5381 | 5187 | 5187 |
| Control mean | 0.41 | 0.92 | 0.38 | 0.93 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## Appendix F. Accounting for multiple hypothesis testing

Figure F1: Correcting p-values following Anderson (2008)



## Appendix G. Heterogeneity - main results

Table G1: Heterogeneity for short-term outcomes by patriarchal norms only for females - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | 0.001 | 0.031 | -0.008 | 0.090*** | 0.004 | 0.048 | 0.006 |
| | (0.026) | (0.025) | (0.009) | (0.024) | (0.028) | (0.029) | (0.014) |
| CEP X Patriarchal | 0.038 | 0.015 | 0.018 | -0.033 | 0.050 | -0.052 | -0.005 |
| | (0.030) | (0.027) | (0.011) | (0.028) | (0.031) | (0.032) | (0.016) |
| Patriarchal | -0.039* | 0.034* | -0.020** | 0.030 | -0.032 | 0.048** | 0.016* |
| | (0.020) | (0.018) | (0.008) | (0.022) | (0.022) | (0.022) | (0.008) |
| N | 4794 | 4794 | 4794 | 4794 | 4794 | 4794 | 4794 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - Patriarchal | 0.040 | 0.046 | 0.010 | 0.057 | 0.054 | -0.004 | 0.000 |
| P-value | 0.038 | 0.014 | 0.207 | 0.011 | 0.011 | 0.827 | 0.968 |
| ITT2 - Non-patriarchal | 0.001 | 0.031 | -0.008 | 0.090 | 0.004 | 0.048 | 0.006 |
| P-value | 0.962 | 0.211 | 0.381 | 0.001 | 0.886 | 0.102 | 0.672 |
| ITT1-ITT2 | 0.038 | 0.015 | 0.018 | -0.033 | 0.050 | -0.052 | -0.005 |
| P-value | 0.203 | 0.579 | 0.112 | 0.249 | 0.121 | 0.112 | 0.726 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table G2: Heterogeneity for medium-term outcomes by patriarchal norms only for females - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course comple-ted (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | -0.014 | 0.014 | 0.014 | 0.023* | 0.025 | -0.006 | -0.022 |
| | (0.018) | (0.026) | (0.030) | (0.013) | (0.027) | (0.009) | (0.019) |
| CEP X Patriarchal | -0.012 | -0.005 | 0.026 | -0.021 | -0.022 | 0.001 | -0.023 |
| | (0.025) | (0.022) | (0.038) | (0.016) | (0.032) | (0.012) | (0.022) |
| Patriarchal | 0.006 | 0.031* | -0.052* | -0.027** | 0.020 | 0.001 | 0.021 |
| | (0.016) | (0.016) | (0.028) | (0.013) | (0.020) | (0.009) | (0.014) |
| N | 4658 | 4658 | 1459 | 4658 | 4605 | 4605 | 4658 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - Patriarchal | -0.026 | 0.009 | 0.040 | 0.002 | 0.002 | -0.005 | -0.045 |
| P-value | 0.176 | 0.709 | 0.053 | 0.839 | 0.902 | 0.517 | 0.015 |
| ITT2 - Non-patriarchal | -0.014 | 0.014 | 0.014 | 0.023 | 0.025 | -0.006 | -0.022 |
| P-value | 0.429 | 0.594 | 0.634 | 0.077 | 0.364 | 0.509 | 0.250 |
| ITT1-ITT2 | -0.012 | -0.005 | 0.026 | -0.021 | -0.022 | 0.001 | -0.023 |
| P-value | 0.635 | 0.820 | 0.496 | 0.195 | 0.492 | 0.936 | 0.293 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table G3: Heterogeneity for short-term outcomes by mobility - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | 0.008 | 0.020 | 0.001 | 0.055*** | 0.042** | 0.003 | -0.003 |
| | (0.020) | (0.018) | (0.008) | (0.020) | (0.020) | (0.015) | (0.009) |
| CEP X More mobile | 0.032 | 0.045* | 0.002 | 0.044 | -0.024 | 0.034* | 0.016 |
| | (0.026) | (0.024) | (0.009) | (0.031) | (0.025) | (0.020) | (0.015) |
| More mobile | 0.011 | -0.028 | 0.008 | 0.028 | 0.040** | 0.017 | 0.002 |
| | (0.016) | (0.017) | (0.006) | (0.022) | (0.017) | (0.013) | (0.011) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - More mobile | 0.040 | 0.065 | 0.003 | 0.099 | 0.018 | 0.037 | 0.014 |
| P-value | 0.041 | 0.007 | 0.614 | 0.000 | 0.407 | 0.110 | 0.274 |
| ITT2 - Less mobile | 0.008 | 0.020 | 0.001 | 0.055 | 0.042 | 0.003 | -0.003 |
| P-value | 0.695 | 0.269 | 0.930 | 0.009 | 0.044 | 0.839 | 0.775 |
| ITT1-ITT2 | 0.032 | 0.045 | 0.002 | 0.044 | -0.024 | 0.034 | 0.016 |
| P-value | 0.220 | 0.069 | 0.794 | 0.167 | 0.346 | 0.095 | 0.284 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table G4: Heterogeneity for medium-term outcomes by mobility - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course comple-ted (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | -0.009 | -0.008 | -0.007 | 0.008 | -0.008 | -0.005 | -0.028* |
| | (0.021) | (0.020) | (0.023) | (0.011) | (0.019) | (0.007) | (0.015) |
| CEP X More mobile | 0.003 | 0.037 | 0.037 | -0.011 | 0.049 | 0.009 | -0.008 |
| | (0.028) | (0.026) | (0.044) | (0.019) | (0.031) | (0.011) | (0.022) |
| More mobile | -0.004 | -0.019 | -0.019 | 0.018 | 0.002 | 0.005 | -0.007 |
| | (0.019) | (0.020) | (0.036) | (0.013) | (0.023) | (0.007) | (0.017) |
| N | 5244 | 5244 | 1652 | 5244 | 5187 | 5187 | 5244 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - More mobile | -0.006 | 0.029 | 0.030 | -0.003 | 0.042 | 0.004 | -0.036 |
| P-value | 0.723 | 0.307 | 0.350 | 0.847 | 0.081 | 0.635 | 0.080 |
| ITT2 - Less mobile | -0.009 | -0.008 | -0.007 | 0.008 | -0.008 | -0.005 | -0.028 |
| P-value | 0.657 | 0.712 | 0.750 | 0.494 | 0.694 | 0.420 | 0.073 |
| ITT1-ITT2 | 0.003 | 0.037 | 0.037 | -0.011 | 0.049 | 0.009 | -0.008 |
| P-value | 0.912 | 0.160 | 0.398 | 0.589 | 0.123 | 0.385 | 0.712 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table G5: Heterogeneity for short-term outcomes by meritocracy - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | 0.001 | 0.006 | -0.011 | 0.041* | 0.037 | -0.002 | 0.020 |
| | (0.025) | (0.025) | (0.009) | (0.022) | (0.029) | (0.020) | (0.013) |
| CEP X Meritocratic | 0.028 | 0.048 | 0.018** | 0.045* | -0.007 | 0.027 | -0.025* |
| | (0.025) | (0.035) | (0.009) | (0.026) | (0.031) | (0.026) | (0.014) |
| Meritocratic | 0.009 | 0.015 | -0.001 | 0.014 | 0.003 | 0.019 | 0.016* |
| | (0.019) | (0.018) | (0.006) | (0.020) | (0.022) | (0.018) | (0.009) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - Meritocratic | 0.029 | 0.054 | 0.008 | 0.086 | 0.030 | 0.025 | -0.005 |
| P-value | 0.052 | 0.019 | 0.213 | 0.000 | 0.090 | 0.213 | 0.578 |
| ITT2 - Non-meritocratic | 0.001 | 0.006 | -0.011 | 0.041 | 0.037 | -0.002 | 0.020 |
| P-value | 0.956 | 0.821 | 0.222 | 0.064 | 0.211 | 0.908 | 0.120 |
| ITT1-ITT2 | 0.028 | 0.048 | 0.018 | 0.045 | -0.007 | 0.027 | -0.025 |
| P-value | 0.269 | 0.178 | 0.036 | 0.087 | 0.817 | 0.291 | 0.082 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table G6: Heterogeneity for medium-term outcomes by meritocracy - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | -0.020 | -0.008 | -0.013 | 0.014 | 0.031 | 0.004 | -0.026 |
| | (0.021) | (0.028) | (0.030) | (0.014) | (0.026) | (0.010) | (0.021) |
| CEP X Meritocratic | 0.018 | 0.022 | 0.031 | -0.015 | -0.029 | -0.008 | -0.008 |
| | (0.030) | (0.030) | (0.042) | (0.016) | (0.027) | (0.012) | (0.030) |
| Meritocratic | -0.016 | -0.030 | -0.007 | 0.005 | 0.034* | 0.012 | -0.008 |
| | (0.020) | (0.022) | (0.032) | (0.010) | (0.017) | (0.010) | (0.020) |
| N | 5244 | 5244 | 1652 | 5244 | 5187 | 5187 | 5244 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - Meritocratic | -0.002 | 0.014 | 0.017 | -0.002 | 0.002 | -0.004 | -0.034 |
| P-value | 0.917 | 0.541 | 0.438 | 0.866 | 0.882 | 0.461 | 0.072 |
| ITT2 - Non-meritocratic | -0.020 | -0.008 | -0.013 | 0.014 | 0.031 | 0.004 | -0.026 |
| P-value | 0.333 | 0.765 | 0.663 | 0.323 | 0.238 | 0.699 | 0.233 |
| ITT1-ITT2 | 0.018 | 0.022 | 0.031 | -0.015 | -0.029 | -0.008 | -0.008 |
| P-value | 0.536 | 0.454 | 0.465 | 0.343 | 0.294 | 0.487 | 0.790 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

## Appendix H. Heterogeneity - other results

Table H1: Heterogeneity for short-term outcomes by school type - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | -0.018 | -0.016 | 0.007 | 0.066*** | 0.079** | 0.019 | -0.010 |
| | (0.037) | (0.029) | (0.010) | (0.018) | (0.035) | (0.030) | (0.012) |
| CEP X Same-sex | 0.044 | 0.063* | -0.006 | 0.006 | -0.056 | -0.005 | 0.016 |
| | (0.042) | (0.033) | (0.012) | (0.028) | (0.040) | (0.035) | (0.015) |
| Same-sex | 0.001 | -0.025 | 0.003 | 0.007 | 0.034** | -0.018 | -0.002 |
| | (0.029) | (0.022) | (0.006) | (0.016) | (0.016) | (0.019) | (0.014) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - Same-sex | 0.027 | 0.047 | 0.001 | 0.072 | 0.024 | 0.015 | 0.006 |
| P-value | 0.116 | 0.010 | 0.939 | 0.000 | 0.208 | 0.407 | 0.501 |
| ITT2 - Coed | -0.018 | -0.016 | 0.007 | 0.066 | 0.079 | 0.019 | -0.010 |
| P-value | 0.640 | 0.589 | 0.513 | 0.001 | 0.029 | 0.522 | 0.393 |
| ITT1-ITT2 | 0.044 | 0.063 | -0.006 | 0.006 | -0.056 | -0.005 | 0.016 |
| P-value | 0.299 | 0.065 | 0.597 | 0.829 | 0.171 | 0.896 | 0.293 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H2: Heterogeneity for medium-term outcomes by school type - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course comple-ted (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | 0.069*** | 0.013 | -0.070** | -0.020 | 0.031 | -0.013 | -0.032 |
| | (0.017) | (0.039) | (0.033) | (0.022) | (0.035) | (0.012) | (0.021) |
| CEP X Same-sex | -0.091*** | -0.007 | 0.092** | 0.028 | -0.024 | 0.013 | 0.001 |
| | (0.022) | (0.044) | (0.039) | (0.024) | (0.038) | (0.013) | (0.027) |
| Same-sex | 0.018 | 0.018 | 0.046*** | -0.020 | -0.052*** | 0.000 | 0.039*** |
| | (0.019) | (0.043) | (0.016) | (0.018) | (0.019) | (0.010) | (0.014) |
| N | 5244 | 5244 | 1652 | 5244 | 5187 | 5187 | 5244 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - Same-sex | -0.023 | 0.006 | 0.022 | 0.008 | 0.008 | 0.000 | -0.031 |
| P-value | 0.117 | 0.804 | 0.218 | 0.397 | 0.613 | 0.933 | 0.055 |
| ITT2 - Coed | 0.069 | 0.013 | -0.070 | -0.020 | 0.031 | -0.013 | -0.032 |
| P-value | 0.000 | 0.747 | 0.041 | 0.376 | 0.376 | 0.275 | 0.148 |
| ITT1-ITT2 | -0.091 | -0.007 | 0.092 | 0.028 | -0.024 | 0.013 | 0.001 |
| P-value | 0.000 | 0.877 | 0.022 | 0.260 | 0.534 | 0.305 | 0.977 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H3: Heterogeneity for short-term outcomes by stream - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | 0.026 | 0.052** | -0.003 | 0.034 | 0.034 | 0.024 | 0.006 |
| | (0.033) | (0.022) | (0.007) | (0.023) | (0.024) | (0.024) | (0.014) |
| CEP X Arts | -0.010 | -0.021 | 0.007 | 0.054* | -0.002 | -0.013 | -0.003 |
| | (0.039) | (0.026) | (0.010) | (0.028) | (0.024) | (0.025) | (0.016) |
| Arts | 0.086* | -0.207*** | -0.016* | -0.115*** | -0.172*** | -0.087*** | 0.007 |
| | (0.043) | (0.019) | (0.008) | (0.024) | (0.023) | (0.027) | (0.015) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - Arts | 0.017 | 0.031 | 0.004 | 0.087 | 0.032 | 0.012 | 0.002 |
| P-value | 0.348 | 0.114 | 0.650 | 0.000 | 0.082 | 0.503 | 0.779 |
| ITT2 - Non-arts | 0.026 | 0.052 | -0.003 | 0.034 | 0.034 | 0.024 | 0.006 |
| P-value | 0.421 | 0.024 | 0.607 | 0.153 | 0.161 | 0.308 | 0.680 |
| ITT1-ITT2 | -0.010 | -0.021 | 0.007 | 0.054 | -0.002 | -0.013 | -0.003 |
| P-value | 0.805 | 0.427 | 0.500 | 0.063 | 0.931 | 0.615 | 0.835 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H4: Heterogeneity for medium-term outcomes by stream - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | 0.017 | 0.030 | -0.020 | -0.005 | 0.028 | 0.008 | -0.023 |
| | (0.026) | (0.035) | (0.029) | (0.014) | (0.030) | (0.009) | (0.018) |
| CEP X Arts | -0.035 | -0.033 | 0.038 | 0.012 | -0.024 | -0.014 | -0.011 |
| | (0.029) | (0.035) | (0.038) | (0.018) | (0.035) | (0.011) | (0.026) |
| Arts | 0.230*** | -0.384*** | 0.151*** | 0.075*** | -0.102*** | -0.002 | 0.086*** |
| | (0.027) | (0.029) | (0.025) | (0.016) | (0.034) | (0.011) | (0.020) |
| N | 5244 | 5244 | 1652 | 5244 | 5187 | 5187 | 5244 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - Arts | -0.019 | -0.003 | 0.018 | 0.007 | 0.005 | -0.006 | -0.034 |
| P-value | 0.236 | 0.892 | 0.392 | 0.527 | 0.782 | 0.358 | 0.062 |
| ITT2 - Non-arts | 0.017 | 0.030 | -0.020 | -0.005 | 0.028 | 0.008 | -0.023 |
| P-value | 0.518 | 0.398 | 0.484 | 0.741 | 0.354 | 0.364 | 0.202 |
| ITT1-ITT2 | -0.035 | -0.033 | 0.038 | 0.012 | -0.024 | -0.014 | -0.011 |
| P-value | 0.237 | 0.360 | 0.312 | 0.526 | 0.500 | 0.217 | 0.677 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H5: Heterogeneity for short-term outcomes by gender - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | -0.022 | 0.040 | -0.017 | 0.060 | 0.010 | 0.011 | 0.012 |
| | (0.042) | (0.048) | (0.015) | (0.048) | (0.037) | (0.030) | (0.021) |
| CEP X Female | 0.047 | -0.003 | 0.021 | 0.013 | 0.025 | 0.006 | -0.010 |
| | (0.046) | (0.050) | (0.016) | (0.055) | (0.041) | (0.033) | (0.023) |
| Female | 0.029 | 0.052 | -0.004 | 0.118*** | 0.003 | 0.016 | -0.010 |
| | (0.030) | (0.043) | (0.010) | (0.022) | (0.028) | (0.023) | (0.021) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - Female | 0.025 | 0.037 | 0.004 | 0.073 | 0.035 | 0.016 | 0.002 |
| P-value | 0.123 | 0.035 | 0.538 | 0.000 | 0.055 | 0.347 | 0.777 |
| ITT2 - Male | -0.022 | 0.040 | -0.017 | 0.060 | 0.010 | 0.011 | 0.012 |
| P-value | 0.604 | 0.411 | 0.249 | 0.222 | 0.787 | 0.722 | 0.563 |
| ITT1-ITT2 | 0.047 | -0.003 | 0.021 | 0.013 | 0.025 | 0.006 | -0.010 |
| P-value | 0.307 | 0.956 | 0.190 | 0.821 | 0.545 | 0.869 | 0.659 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H6: Heterogeneity for medium-term outcomes by gender - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | 0.061 | -0.016 | -0.149*** | -0.056* | 0.020 | 0.018 | 0.004 |
| | (0.051) | (0.039) | (0.055) | (0.029) | (0.031) | (0.019) | (0.023) |
| CEP X Female | -0.078 | 0.026 | 0.177*** | 0.067** | -0.009 | -0.022 | -0.040 |
| | (0.053) | (0.045) | (0.060) | (0.032) | (0.032) | (0.020) | (0.030) |
| Female | 0.055** | -0.077* | -0.127** | -0.122*** | -0.010 | 0.001 | 0.235*** |
| | (0.023) | (0.039) | (0.054) | (0.031) | (0.022) | (0.015) | (0.027) |
| N | 5244 | 5244 | 1652 | 5244 | 5187 | 5187 | 5244 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - Female | -0.017 | 0.010 | 0.028 | 0.011 | 0.011 | -0.004 | -0.036 |
| P-value | 0.215 | 0.658 | 0.084 | 0.201 | 0.485 | 0.434 | 0.022 |
| ITT2 - Male | 0.061 | -0.016 | -0.149 | -0.056 | 0.020 | 0.018 | 0.004 |
| P-value | 0.237 | 0.678 | 0.009 | 0.063 | 0.536 | 0.354 | 0.852 |
| ITT1-ITT2 | -0.078 | 0.026 | 0.177 | 0.067 | -0.009 | -0.022 | -0.040 |
| P-value | 0.149 | 0.558 | 0.005 | 0.039 | 0.781 | 0.282 | 0.181 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H7: Heterogeneity for short-term outcomes by caste - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | 0.003 | 0.044* | -0.001 | 0.029 | 0.022 | -0.007 | 0.006 |
| | (0.028) | (0.026) | (0.012) | (0.031) | (0.030) | (0.029) | (0.018) |
| CEP X Upper caste | 0.028 | 0.004 | 0.008 | 0.070* | 0.015 | 0.033 | -0.002 |
| | (0.028) | (0.033) | (0.012) | (0.036) | (0.029) | (0.030) | (0.019) |
| Upper caste | -0.048** | -0.019 | -0.006 | -0.042 | 0.005 | -0.011 | 0.006 |
| | (0.021) | (0.025) | (0.008) | (0.028) | (0.020) | (0.017) | (0.016) |
| N | 4241 | 4241 | 4241 | 4241 | 4241 | 4241 | 4241 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - Upper caste | 0.031 | 0.048 | 0.007 | 0.099 | 0.037 | 0.026 | 0.004 |
| P-value | 0.057 | 0.027 | 0.373 | 0.000 | 0.046 | 0.131 | 0.629 |
| ITT2 - Non-upper caste | 0.003 | 0.044 | -0.001 | 0.029 | 0.022 | -0.007 | 0.006 |
| P-value | 0.922 | 0.098 | 0.911 | 0.353 | 0.462 | 0.823 | 0.740 |
| ITT1-ITT2 | 0.028 | 0.004 | 0.008 | 0.070 | 0.015 | 0.033 | -0.002 |
| P-value | 0.324 | 0.907 | 0.527 | 0.056 | 0.610 | 0.285 | 0.931 |

*Notes:* All specifications include baseline covariates. The information for caste was asked only in the follow up surveys since it was not allowed to ask in school by school authorities and 904 students who participated in the followup survey did not participate in the endline survey. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table H8: Heterogeneity for medium-term outcomes by caste - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course comple-ted (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | -0.023 | -0.009 | -0.001 | 0.007 | -0.016 | -0.007 | -0.013 |
| | (0.022) | (0.029) | (0.028) | (0.015) | (0.026) | (0.013) | (0.025) |
| CEP X Upper caste | 0.024 | 0.031 | 0.011 | -0.003 | 0.037 | 0.009 | -0.028 |
| | (0.031) | (0.034) | (0.031) | (0.017) | (0.026) | (0.014) | (0.032) |
| Upper caste | -0.036 | -0.031 | -0.003 | 0.004 | -0.005 | 0.014 | 0.038 |
| | (0.023) | (0.025) | (0.017) | (0.013) | (0.016) | (0.009) | (0.024) |
| N | 5123 | 5123 | 1612 | 5123 | 5085 | 5085 | 5123 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - Upper caste | 0.001 | 0.022 | 0.010 | 0.004 | 0.021 | 0.003 | -0.041 |
| P-value | 0.943 | 0.370 | 0.566 | 0.732 | 0.143 | 0.674 | 0.023 |
| ITT2 - Non-upper caste | -0.023 | -0.009 | -0.001 | 0.007 | -0.016 | -0.007 | -0.013 |
| P-value | 0.313 | 0.751 | 0.980 | 0.643 | 0.553 | 0.590 | 0.614 |
| ITT1-ITT2 | 0.024 | 0.031 | 0.011 | -0.003 | 0.037 | 0.009 | -0.028 |
| P-value | 0.439 | 0.355 | 0.723 | 0.851 | 0.169 | 0.510 | 0.388 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H9: Heterogeneity for short-term outcomes by religion - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | 0.037 | 0.058 | 0.010 | 0.110*** | 0.015 | 0.037 | 0.012 |
| | (0.035) | (0.038) | (0.021) | (0.029) | (0.032) | (0.039) | (0.014) |
| CEP X Hindu | -0.015 | -0.012 | -0.005 | -0.040 | 0.018 | -0.024 | -0.010 |
| | (0.039) | (0.039) | (0.022) | (0.030) | (0.038) | (0.041) | (0.017) |
| Hindu | 0.037 | 0.007 | 0.025 | 0.098*** | 0.039* | 0.016 | -0.004 |
| | (0.032) | (0.029) | (0.015) | (0.023) | (0.023) | (0.029) | (0.010) |
| N | 4305 | 4305 | 4305 | 4305 | 4305 | 4305 | 4305 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - Hindu | 0.022 | 0.046 | 0.005 | 0.070 | 0.033 | 0.013 | 0.002 |
| P-value | 0.205 | 0.021 | 0.419 | 0.001 | 0.104 | 0.473 | 0.808 |
| ITT2 - Non-hindu | 0.037 | 0.058 | 0.010 | 0.110 | 0.015 | 0.037 | 0.012 |
| P-value | 0.297 | 0.133 | 0.654 | 0.000 | 0.636 | 0.345 | 0.392 |
| ITT1-ITT2 | -0.015 | -0.012 | -0.005 | -0.040 | 0.018 | -0.024 | -0.010 |
| P-value | 0.701 | 0.760 | 0.838 | 0.185 | 0.640 | 0.557 | 0.557 |

*Notes:* All specifications include baseline covariates. The information for religion was asked only in the follow up surveys since it was not allowed to ask in school by school authorities and 904 students who participated in the followup survey did not participate in the endline survey. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table H10: Heterogeneity for medium-term outcomes by religion - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | -0.008 | -0.020 | 0.018 | 0.044** | -0.024 | -0.005 | -0.066*** |
| | (0.032) | (0.039) | (0.035) | (0.019) | (0.028) | (0.014) | (0.024) |
| CEP X Hindu | 0.003 | 0.038 | -0.017 | -0.051** | 0.046 | 0.005 | 0.043 |
| | (0.035) | (0.037) | (0.042) | (0.023) | (0.030) | (0.016) | (0.026) |
| Hindu | 0.069** | -0.006 | 0.017 | 0.024** | -0.019 | 0.006 | -0.110*** |
| | (0.026) | (0.031) | (0.033) | (0.010) | (0.019) | (0.013) | (0.018) |
| N | 5194 | 5194 | 1633 | 5194 | 5158 | 5158 | 5194 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - Hindu | -0.006 | 0.018 | 0.001 | -0.006 | 0.022 | 0.000 | -0.023 |
| P-value | 0.708 | 0.387 | 0.951 | 0.557 | 0.142 | 0.992 | 0.112 |
| ITT2 - Non-hindu | -0.008 | -0.020 | 0.018 | 0.044 | -0.024 | -0.005 | -0.066 |
| P-value | 0.799 | 0.608 | 0.616 | 0.022 | 0.394 | 0.747 | 0.010 |
| ITT1-ITT2 | 0.003 | 0.038 | -0.017 | -0.051 | 0.046 | 0.005 | 0.043 |
| P-value | 0.943 | 0.310 | 0.698 | 0.033 | 0.126 | 0.772 | 0.108 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H11: Heterogeneity for short-term outcomes by grades - ITT controlling for baseline covariates

| | (1) Enroll in course (H1) | (2) Entrance exam (H2) | (3) Work for pay (H3) | (4) Interest based job (H4) | (5) Intergen. mobility (H5) | (6) Locus of control (H10) | (7) Self-efficacy (H11) |
|---|---|---|---|---|---|---|---|
| CEP | 0.011 | 0.011 | 0.003 | 0.115*** | 0.001 | -0.005 | 0.007 |
| | (0.021) | (0.021) | (0.011) | (0.019) | (0.022) | (0.019) | (0.011) |
| CEP X High performer | 0.022 | 0.063** | -0.001 | -0.080*** | 0.059** | 0.042 | -0.005 |
| | (0.028) | (0.023) | (0.011) | (0.026) | (0.026) | (0.025) | (0.014) |
| High performer | -0.021 | 0.052*** | 0.020*** | 0.147*** | 0.082*** | 0.041** | 0.014 |
| | (0.018) | (0.014) | (0.006) | (0.020) | (0.019) | (0.020) | (0.010) |
| N | 5048 | 5048 | 5048 | 5048 | 5048 | 5048 | 5048 |
| Control mean | 0.513 | 0.574 | 0.965 | 0.439 | 0.534 | 0.406 | 0.925 |
| ITT1 - High performer | 0.032 | 0.073 | 0.002 | 0.036 | 0.060 | 0.037 | 0.002 |
| P-value | 0.106 | 0.001 | 0.626 | 0.080 | 0.005 | 0.083 | 0.846 |
| ITT2 - Low performer | 0.011 | 0.011 | 0.003 | 0.115 | 0.001 | -0.005 | 0.007 |
| P-value | 0.612 | 0.620 | 0.753 | 0.000 | 0.957 | 0.794 | 0.523 |
| ITT1-ITT2 | 0.022 | 0.063 | -0.001 | -0.080 | 0.059 | 0.042 | -0.005 |
| P-value | 0.441 | 0.011 | 0.932 | 0.004 | 0.027 | 0.104 | 0.701 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

Table H12: Heterogeneity for medium-term outcomes by grades - ITT controlling for baseline covariates

| | (1) Enrolled in course (H6) | (2) Took entrance exam (H7) | (3) Course completed (H8) | (4) Working for pay (H9) | (5) Locus of control (H12) | (6) Self-efficacy (H13) | (7) Spending time in household work (Not pre-registered) |
|---|---|---|---|---|---|---|---|
| CEP | -0.029 | -0.010 | -0.015 | 0.003 | -0.003 | -0.004 | -0.034* |
| | (0.020) | (0.021) | (0.022) | (0.015) | (0.020) | (0.007) | (0.020) |
| CEP X High performer | 0.048* | 0.037 | 0.052 | 0.003 | 0.033 | 0.004 | -0.003 |
| | (0.027) | (0.029) | (0.038) | (0.021) | (0.025) | (0.010) | (0.026) |
| High performer | 0.079*** | 0.132*** | -0.025 | -0.026** | 0.051*** | 0.010 | -0.102*** |
| | (0.019) | (0.019) | (0.032) | (0.011) | (0.017) | (0.007) | (0.021) |
| N | 5087 | 5087 | 1597 | 5087 | 5041 | 5041 | 5087 |
| Control mean | 0.698 | 0.293 | 0.213 | 0.099 | 0.385 | 0.956 | 0.320 |
| ITT1 - High performer | 0.018 | 0.027 | 0.037 | 0.005 | 0.029 | 0.000 | -0.037 |
| P-value | 0.366 | 0.307 | 0.185 | 0.649 | 0.121 | 0.975 | 0.030 |
| ITT2 - Low performer | -0.029 | -0.010 | -0.015 | 0.003 | -0.003 | -0.004 | -0.034 |
| P-value | 0.151 | 0.639 | 0.494 | 0.868 | 0.871 | 0.573 | 0.093 |
| ITT1-ITT2 | 0.048 | 0.037 | 0.052 | 0.003 | 0.033 | 0.004 | -0.003 |
| P-value | 0.090 | 0.207 | 0.180 | 0.904 | 0.201 | 0.689 | 0.901 |

*Notes:* All specifications include baseline covariates. Standard errors are clustered at the school level.

## Appendix I. Impact on students' plans and actions

Table I1: 1st and 2nd plans in endline surveys - ITT and LATE accounting for students' attendance - with baseline covariates

| | Course | | Exam | | Work | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | 1st | 2nd | 1st | 2nd | 1st | 2nd | No 2nd |
| CEP | 0.012 | 0.010 | 0.017 | 0.024** | -0.005 | 0.003 | -0.038** |
| | (0.014) | (0.012) | (0.016) | (0.012) | (0.008) | (0.012) | (0.015) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 |
| ***Panel B: LATE*** | | | | | | | |
| Regular CEP student | 0.025 | 0.021 | 0.035 | 0.051** | -0.010 | 0.006 | -0.079** |
| | (0.028) | (0.025) | (0.033) | (0.024) | (0.017) | (0.025) | (0.034) |
| N | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 | 5381 |
| Control mean | 0.31 | 0.21 | 0.43 | 0.20 | 0.11 | 0.17 | 0.39 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table I2: 1st and 2nd actions in follow-up surveys - ITT and LATE accounting for students' attendance - with baseline covariates

| | Course | | Exam | | Work | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | 1st | 2nd | 1st | 2nd | 1st | 2nd | No 2nd |
| CEP | 0.021 | -0.018 | 0.002 | 0.007 | 0.009 | -0.003 | 0.005 |
| | (0.014) | (0.014) | (0.009) | (0.006) | (0.009) | (0.005) | (0.013) |
| N | 5185 | 5185 | 5185 | 5185 | 5185 | 5185 | 5185 |
| ***Panel B: LATE*** | | | | | | | |
| Regular CEP student | 0.045 | -0.039 | 0.005 | 0.014 | 0.019 | -0.006 | 0.011 |
| | (0.031) | (0.031) | (0.019) | (0.013) | (0.019) | (0.011) | (0.028) |
| N | 5185 | 5185 | 5185 | 5185 | 5185 | 5185 | 5185 |
| Control mean | 0.54 | 0.27 | 0.15 | 0.07 | 0.10 | 0.05 | 0.47 |

*Notes:* As pre-registered, all specifications include baseline covariates. Standard errors are clustered at the school level. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

## Appendix J. Mediation

Table J1: Treatment effects on mediators and mediators' effects on outcomes

| | Treatment | Entrance | | Interest-based | | Inter-gen | | Housework | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) Single | (3) Full | (4) Single | (5) Full | (6) Single | (7) Full | (8) Single | (9) Full |
| Information | 0.124*** | 0.130*** | 0.099*** | 0.061*** | 0.053*** | 0.061*** | 0.050*** | 0.061*** | 0.051*** |
| | (0.021) | (0.014) | (0.015) | (0.014) | (0.014) | (0.014) | (0.015) | (0.014) | (0.014) |
| Educational aspiration | -0.008 | 0.208*** | 0.179*** | 0.085*** | 0.076*** | 0.090*** | 0.077*** | 0.097*** | 0.083*** |
| | (0.015) | (0.012) | (0.012) | (0.015) | (0.015) | (0.015) | (0.016) | (0.018) | (0.017) |
| Salary aspiration | 0.019 | 0.128*** | 0.100*** | 0.030** | 0.017 | 0.042*** | 0.030** | 0.044*** | 0.028** |
| | (0.019) | (0.014) | (0.013) | (0.015) | (0.014) | (0.012) | (0.012) | (0.011) | (0.011) |
| Peer communication | -0.008 | 0.070*** | 0.056*** | 0.032** | 0.026* | 0.040*** | 0.033** | 0.050*** | 0.046** |
| | (0.014) | (0.013) | (0.013) | (0.013) | (0.014) | (0.015) | (0.016) | (0.018) | (0.018) |
| Peer influence | -0.092 | 0.014*** | 0.007 | 0.008 | 0.005 | 0.010* | 0.007 | | |
| | (0.116) | (0.005) | (0.005) | (0.005) | (0.005) | (0.006) | (0.006) | | |
| Peer influence - follow-up | -0.256** | | | | | | | 0.015*** | 0.013** |
| | (0.109) | | | | | | | (0.005) | (0.005) |
| Community support | 0.018 | 0.068*** | 0.058*** | -0.029 | -0.032 | 0.032* | 0.028 | 0.018 | 0.014 |
| | (0.011) | (0.017) | (0.016) | (0.021) | (0.021) | (0.018) | (0.018) | (0.015) | (0.014) |
| Facilitators | 0.132*** | 0.026 | 0.033 | -0.011 | -0.008 | -0.010 | -0.007 | -0.033 | -0.024 |
| | (0.012) | (0.026) | (0.024) | (0.021) | (0.021) | (0.022) | (0.024) | (0.021) | (0.019) |
| N | | | 5352 | | 5352 | | 5352 | | 4321 |

*Notes:* As pre-registered, all specifications include baseline covariates. Column 1 shows the treatment's impact on the mediator. The columns 2, 4, and 6 show the effects of the mediators on the outcome variables. The columns 3, 5, and 7 show the full model with all mediators together in a single specification. Standard errors are clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

## References


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. Journal of the American statistical Association 103, 1481–1495.

Baker, R., Bettinger, E., Jacob, B., Marinescu, I., 2018. The effect of labor market information on community college students' major choice. Economics of Education Review 65, 18–30.

Bettinger, E.P., Evans, B.J., 2019. College guidance for all: A randomized experiment in pre-college advising. Journal of Policy Analysis and Management 38, 579–599.

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

Carlsson, M., 2011. Does hiring discrimination cause gender segregation in the swedish labor market? Feminist Economics 17, 71–102.

Carrell, S., Sacerdote, B., 2017. Why do college-going interventions work? American Economic Journal: Applied Economics 9, 124–151.

Chowdhury, A.R., Areias, A.C., Imaizumi, S., Nomura, S., Yamauchi, F., 2018. Reflections of employers' gender preferences in job ads in india: An analysis of online job portal data. World Bank Policy Research Working Paper 8379 .

Deshpande, A., Goel, D., Khanna, S., 2018. Bad karma or discrimination? male–female wage gaps among salaried workers in india. World Development 102, 331–344.

Dex, S., Joshi, H., Macran, S., McCulloch, A., 1998. Women's employment transitions around childbearing. Oxford Bulletin of Economics and Statistics 60, 79–98.

Duflo, E., 2012. Women empowerment and economic development. Journal of Economic Literature 50, 1051–1079.

Duraisamy, M., Duraisamy, P., 2016. Gender wage gap across the wage distribution in different segments of the indian labour market, 1983–2012: exploring the glass ceiling or sticky floor phenomenon. Applied Economics 48, 4098–4111.

Fletcher, E., Pande, R., Moore, C.M.T., 2017. Women and work in india: Descriptive evidence and a review of potential policies. India Policy Forum .

Gehrke, E., Lenel, F., Schupp, C., 2023. Occupational aspirations and investments in education: Experimental evidence from Cambodia. Technical Report. CESifo Working Paper.

Genicot, G., Ray, D., 2017. Aspirations and inequality. Econometrica 85, 489–519.

Genicot, G., Ray, D., 2020. Aspirations and economic behavior. Annual Review of Economics 12, 715–746.

Heckman, J., Pinto, R., Savelyev, P., 2013. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. American Economic Review 103, 2052–2086.

Heckman, J.J., Pinto, R., 2015. Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. Econometric Reviews 34, 6–31.

Imai, K., Keele, L., Yamamoto, T., 2010. Identification, inference and sensitivity analysis for causal mediation effects .

International Institute for Population Sciences and ICF, 2022. National family health survey (nfhs-4), 2015-2016: India.

International Labour Organization, 2022. "ilo modelled estimates and projections database (iloest)" ilostat .

Jayachandran, S., 2020. Social norms as a barrier to women's employment in developing countries. Technical Report. National Bureau of Economic Research.

Jensen, R., 2010. The (perceived) returns to education and the demand for schooling. The Quarterly Journal of Economics 125, 515–548.

Jensen, R., 2012. Do labor market opportunities affect young women's work and family decisions? experimental evidence from india. The Quarterly Journal of Economics 127, 753–792.

Kipchumba, E., Porter, C., Serra, D., Sulaiman, M., 2024. The Impact of Role Models on Youths' Aspirations, Gender Attitudes and Education in Somalia. Technical Report. IZA Discussion Papers.

Klasen, S., 2019. What explains uneven female labor force participation levels and trends in developing countries? The World Bank Research Observer 34, 161–197.

Kuhn, P., Shen, K., 2013. Gender discrimination in job ads: Evidence from china. The Quarterly Journal of Economics 128, 287–336.

Lafortune, J., Riutort, J., Tessada, J., 2018. Role models or individual consulting: The impact of personalizing micro-entrepreneurship training. American Economic Journal: Applied Economics 10, 222–245.

Lee, V.E., Ekstrom, R.B., 1987. Student access to guidance counseling in high school. American Educational Research Journal 24, 287–310.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., Zhang, L., Shi, Y., Chu, J., Rozelle, S., 2013. Can information and counseling help students from poor rural areas go to high school? evidence from china. Journal of Comparative Economics 41, 1012–1025.

Mehta, V., Sai, H., 2021. Freedom of movement: Studying women's mobility in north india. Urbanisation 6, S77–S114.

Ministry of Statistics and Programme Implementation (MOSPI), 2024. Household consumption expenditure survey: 2022–23. URL: https://www.mospi.gov.in/sites/default/files/press_release/HCES_Press_Note_2023-24_27122024_rev.pdf.

Muralidharan, K., 2017. Field experiments in education in developing countries, in: Handbook of economic field experiments. Elsevier. volume 2, pp. 323–385.

National Sample Survey Organization, 2021. India national sample survey 2011-2012 (68th round) - schedule 1.0 (type 1) - consumer expenditure.

Nouman, A., 2022. Two-thirds of delhi govt schools not teaching science in classes 11-12: Rti. The Times of India URL: https://economictimes.indiatimes.com/industry/services/education/two-thirds-of-delhi-govt-schools-not-teaching-science-in-classes-11-12-rti/articleshow/93836914.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.

NSSO, 2023. Annual report periodic labour force surveys (PLFS) - July 2022 - June 2023. Technical Report. National Sample Survey Office, National Statistics Organisation, Ministry of Statistics and Programme Implementation, Government of India.

Porter, C., Serra, D., 2020. Gender differences in the choice of major: The importance of female role models. American Economic Journal: Applied Economics 12, 226–254.

Ray, A., Bala, P.K., Dasgupta, S.A., Srivastava, A., 2020. Understanding the factors influencing career choices in india: from the students' perspectives. International Journal of Indian Culture and Business Management 20, 175–193.

Renée, L., 2025. The long-lerm effects of career guidance in high school and student financial aid: Evidence from a randomized experiment. American Economic Journal: Applied Economics 17, 165–183.

Resnjanskij, S., Ruhose, J., Wiederhold, S., Woessmann, L., Wedel, K., 2024. Can mentoring alleviate family disadvantage in adolescence? a field experiment to improve labor market prospects. Journal of Political Economy 132, 1013–1062.

Riach, P.A., Rich, J., 2006. An experimental investigation of sexual discrimination in hiring in the english labor market. The BE Journal of Economic Analysis & Policy 6, 000010220215380637 1416.

Wiswall, M., Zafar, B., 2015. Determinants of college major choice: Identification using an information experiment. The Review of Economic Studies 82, 791–824.

Zafar, M., 2019. Career guidance in career planning among secondary school students. Asian Journal of Education and Social Studies 5, 1–8.