09 September 2026

To

The Selection Committee
Workshop on Development Economics and Networks for Early-Stage Researchers
National Bureau of Economic Research

**Subject: Application to participate in the workshop**

Dear Committee
I am a 4th-year Economics PhD student at Monash University, Australia, and currently a PhD Fellow at the Harvard Center for International Development. My research focuses on education and human capital formation through randomized field experiments. My experience as a policy practitioner in Bangladesh informs my interest in interventions implementable through existing public institutions.

My PhD research includes two RCTs in India and Bangladesh. The Indian project examines whether adolescent girls' human capital can be improved under household and institutional constraints. Data collection is complete, and we are drafting the working paper. My solo project in Bangladesh estimates the effect of recurring weather events such as rain and radiant heat on school absenteeism among rural students. The study tests a low-cost adaptation to adverse weather conditions.

I am also developing a project to understand why parents who value their children's education still invest very little in the everyday actions that support learning. Through a randomized evaluation, we will test the effectiveness of a scalable intervention to strengthen parental engagement in low-resource settings.

I see the workshop as a great opportunity to learn from other researchers, present my work-in-progress paper, and build research collaborations. I would contribute by sharing my experience implementing school-based experiments in South Asian settings and working with government partners.

My PhD supervisor at Monash University, Professor Liang Choon Wang, can provide a reference regarding my research. He can be contacted at liang.c.wang@monash.edu.

Thank you for considering my application.

Yours sincerely

Imrul Kayes
PhD Candidate in Economics, Monash University &
PhD Fellow, Harvard Center for International Development
Email: imrul_kayes@hks.harvard.edu; Phone: +1 (617) 233-7964

# Imrul Kayes

3 Putnam Gardens, Cambridge, MA 02139, USA
imrul_kayes@hks.harvard.edu; +1 (617) 233-7964


## Education

- Harvard University
  CID PhD Fellow, Harvard Kennedy School (Fall 2026)
- Monash University
  PhD in Economics (Expected completion: August 2027)
- Australian National University
  Master of International and Development Economics
- University of Dhaka
  Bachelor of Business Administration

## PhD Courses Undertaken

1. Microeconomic Theory
2. Macroeconomic Theory
3. Empirical Strategies for Policy Evaluation
4. Field Experiments for Social and Public Policy
5. Research Topics in Advanced Microeconomics
6. Research Topics in Advanced Macroeconomics
7. Macroeconometrics
8. Economic Research Project 1 & 2

## Journal Publication

- Hassan, H., Islam, A., Kayes, I., & Wang, L. C. (2025). Building education resilience through parenting style and out-of-school learning: Field experimental evidence from rural Bangladesh. *Economics of Education Review*, *105*, 102623.

## Research-in-Progress

- *Enhancing educational opportunities: a randomized trial of English and life skills training for underprivileged girls,* with David Blakeslee (NYU-AD), Nishtha Sharma (NYU-AD), Nitin Kumar Bharti (UWA), and Aniruddha Talukdar (LSE).

  In collaboration with the State Education Department and a local NGO, we evaluate a multifaceted educational program for adolescent girls and their parents from backward communities in Karnataka, India.

- *Schooling and Human Capital Formation under Weather Stress: Experimental Evidence on Climate Adaptation from Bangladesh*, as PI

  In partnership with the local government office, I conduct an RCT to assess the impact of a low-cost intervention to reduce the adverse weather-related costs on school attendance in Bangladesh.

- *Fertility choices under housing market pressures: causal evidence from Australia,* with Liang C. Wang.

  We examine the causal effect of housing costs on fertility behavior by applying an instrumental variable approach to longitudinal HILDA data.

## Grants

- USD 250,000; Echidna Giving, with David Blakeslee (NYU), Nitin Bharti (UWA), Nishtha Sharma (NYU), and Aniruddha Talukdar (LSE) (2024)
- AUD 20,000; Monash Department of Economics Research Grant with Liang Choon Wang (2025)
- AUD 1,000; Monash Faculty of Business & Economics Teaching Associate Grant (2025)
- AUD 8,000; Monash Business School Enhanced International Research Experience Grant (2026)
- BDT 700,000; Bangladesh Civil Service Administration Academy Research Grant (2026)

## Scholarships and Awards

- Best Teaching Award, Monash Faculty of Business & Economics (2025)
- Monash Graduate Scholarship (2024-2027)
- Monash International Tuition Scholarship (2024-2027)
- Monash Business School Graduate Research Top-Up Scholarship (2024-2027)
- Monash Business School Dean's Excellence Award (2023)
- Sir John Monash Scholarship for Excellence (2024)
- Australian Development Awards (2013-2014)

## Teaching Experience at Monash University

Taught the following undergraduate and postgraduate units (2023–2025) at Monash University with excellent student evaluations. [Evaluation Reports]

- Principles of Microeconomics
- Intermediate Microeconomics
- Mathematical Economics
- Managerial Economics
- Business in Asia
- Southeast Asian Economies and Global Business

## Journal Referee

- Energy Economics
- Australian Economic Papers

## Public Service Experience

- Deputy Director (July 2020 – July 2022)
  Anti-Corruption Commission, Bangladesh
- Additional Deputy Commissioner (October 2019 – June 2020)
  Narsingdi District, Bangladesh
- National Program Officer (October 2017 – September 2019)
  United Nations Development Program (UNDP), Bangladesh
- Assistant Director (May 2016 – September 2017)
  Bangladesh Civil Service Administration Academy, Bangladesh
- Assistant Commissioner (December 2010 – April 2016)
  Kishoreganj, Khagrachari, and Sunamganj Districts, Bangladesh

## Motivation

My experience as a policy practitioner in Bangladesh has shaped my interest in how public programs can improve education among disadvantaged children. I am particularly interested in the barriers that persist even when schooling is available. The barriers include limited support for learning at home and in schools, as well as disruptions to attendance. My research examines how these constraints affect human capital formation and whether low-cost interventions delivered through existing public institutions can help address them.

I mainly conduct randomized field experiments to study education and human capital formation in South Asia. Across projects, I examine how additional educational support, parental engagement, and protection from weather-related disruptions affect learning outcomes. My recent research idea is the relationship between parents' educational aspirations and their actions. Parents may value education highly yet struggle to translate that commitment into consistent support for their children's learning. I would like to understand what explains this gap and whether scalable interventions can narrow it.

I am keen to attend the workshop as I am currently turning my fieldwork into research papers. I appreciate discussions that will help me link findings from specific settings to broader economic arguments. I expect useful comments on refining the interpretation of treatment effects and evaluating evidence for alternative mechanisms.

The workshop would also offer an opportunity to build research relationships beyond my existing collaborations. I would welcome candid exchanges about both conceptual questions and the practical challenges of field experiments. I could contribute experience working with schools and government partners in South Asia, while learning how other researchers approach similar questions in different settings.

# Human Capital Formation under Constraints: A Randomized Trial among Disadvantaged Girls in India*

Nitin Bharti † David Blakeslee ‡ Amalu Joseph § Imrul Kayes ¶
Nishtha Sharma ‖ Aniruddha Talukdar **

September 4, 2026

**Abstract**

This cluster-randomized controlled trial examines whether human capital of adolescent girls can be improved under persistent household and institutional constraints. The study involves government residential schools for girls from disadvantaged communities in India. Treatment 1 received foundational English and life-skills lessons through hostel wardens. Treatment 2 received the same intervention as Treatment 1, while their parents received weekly audio messages through an Interactive Voice Response (IVR) system. The program improves both cognitive and non-cognitive dimensions of human capital. The combined treatment improves English skills among students in Grades 9–10 by 0.16 SD and *confidence index* by 0.13 SD, while educational aspirations rise marginally. Treatment 2 also increases school attendance and study time, reduces perceived gender and family barriers, and improves agency. Parents report higher importance of girls' education and greater confidence in their daughters' educational and professional success, including higher expected earnings. The findings show that a cost-effective program using existing institutional resources can strengthen human capital even when constraints remain largely unchanged.

**Keywords:** Human capital, girls' education, remedial education, India, RCT.
**JEL codes:** I21, I24, I25, I26, J16, O15.

*This study was supported by a grant from Echidna Giving. We obtained ethics approval from Monash University Human Research Ethics Committee. The trial was pre-registered with AEA RCT Registry (AEARCTR-0015254), *https://www.socialscienceregistry.org/trials/15254*. We thank the State Government of Karnataka, India for administrative support, and SwaTaleem Foundation for designing and implementing the program.

†University of Western Australia, nitin.bharti@uwa.edu.au
‡NYU Abu Dhabi, dsb12@nyu.edu
§josephamalukunnamkottu@gmail.com
¶Monash University; Harvard Center for International Development, imrul.kayes@monash.edu
‖Charles University, nishthasharma.econ@gmail.com
**LSE, a.talukdar@lse.ac.uk

# 1 Introduction

Human capital formation is particularly difficult when children face multiple constraints at the same time. Household and institutional constraints heighten the difficulties for girls from disadvantaged communities. These constraints often interact with gendered expectations about education, work, mobility, and marriage (Duflo, 2012; Jayachandran, 2015; Dhar et al., 2019; Desai and Kulkarni, 2008). Poor foundational literacy, limited aspirations and confidence, restrictive social norms, and limited support from families or institutions diminish their ability to invest in their own education (Heckman and Mosso, 2014; Ray, 2006; La Ferrara, 2019; Jayachandran, 2021). This raises a policy question: how can human capital formation of the disadvantaged girls be facilitated even when many of the underlying constraints remain in place? We study this question through a cluster-randomized trial in special residential schools for disadvantaged girls in India. In collaboration with a local NGO, and the State Education Department, we provide English and life skills lessons to girls in Karnataka. We also engage their parents through Interactive Voice Response (IVR)[1] technology, delivering targeted messages relevant to the girls' education and development.

In India, the government operates the Kasturba Gandhi Balika Vidyalaya (KGBV)[2] system to provide residential education for girls from disadvantaged communities. KGBV hostels offer free schooling, accommodation, and meals. Approximately 5,000 KGBV hostels function across India, with about 220 located in Karnataka. The Karnataka State Government granted us administrative access to the KGBV hostel network, and *SwaTaleem Foundation*, an India-based non-governmental organization, implemented the intervention. *SwaTaleem* had previously delivered English and life-skills programs in KGBV hostels in another Indian state. Expanding a similar program to Karnataka allowed us to conduct this randomized controlled trial.

Several studies provide evidence that targeted interventions can effectively reduce gender-based inequalities in education and employment while promoting more equitable gender norms. Andrew et al. (2022) show that group education and sports sessions in Rajasthan increase girls' school attendance and delay early marriage. Interventions that enhance educational access have also proven effective. Giacobino et al. (2024) evaluate a scholarship program in Niger, finding that financial support reduces dropout rates and child marriage while improving girls' aspirations. Likewise, Bandiera et al. (2020) demonstrate that life skills and vocational training for adolescent girls in Uganda increase their empowerment, reduce early pregnancies, and improve time allocation between school and work. School-based gender sensitization programs in Haryana enhanced adolescents' gender-equal attitudes and

[1] Interactive Voice Response (IVR) is an automated phone system that enables users to access information through a touch-tone keypad or voice command interface.

[2] Throughout this document, we use KGBV and KGBV hostel interchangeably.

behaviors Dhar et al. (2022). Additionally, media-based interventions have effectively shifted societal norms, as seen in Jensen and Oster (2009) and Ferrara et al. (2012), where exposure to television content altered fertility preferences and educational aspirations. These studies collectively highlight that structural barriers for women can be addressed through well-designed educational and social interventions.

The interventions in this study are designed following literature highlighting the role of cognitive and non-cognitive skills in enhancing economic outcomes. English proficiency is an important cognitive skill in India, where it serves as a *lingua franca* in higher education and formal employment. English fluency facilitates access to higher-paying jobs and is essential for pursuing higher education (Jain et al., 2019). In India, fluent English speakers earn 34% higher hourly wages, and those with limited proficiency earn 13% more than non-speakers (Azam et al., 2013). Programs integrating English lessons demonstrate the potential to improve employability and increase wages among low-income individuals (Loyalka et al., 2025).

Non-cognitive traits, such as aspirations, gendered norms, and socio-emotional learning, are vital for future success and can be developed through targeted interventions (Kautz et al., 2014). Studies show that individuals with high aspirations and growth mindsets tend to achieve better educational, economic, and social outcomes, as they are more likely to persist in achieving long-term goals despite challenges (Card, 2001; Heckman and Mosso, 2014; La Ferrara, 2019; Sisk et al., 2018). Strengthening the capacity to aspire can empower the poor to change their conditions of poverty (Ray, 2006), but aspiration failures are prevalent among marginalized populations due to poverty, unequal opportunities, and restrictive social norms (Appadurai, 2007; Dalton et al., 2016). Additionally, addressing restrictive social and gender norms among these communities can encourage female employment in the formal sector and help break the cycle of poverty (Jayachandran, 2021).

Adolescents represent a critical demographic for interventions (Sheehan et al., 2017; Steinberg, 2014). For adolescent girls, especially those from rural regions, this period often coincides with heightened challenges, such as restricted mobility, disproportionate household responsibilities, early marriage, and limited opportunities (Dunn, 1993). Targeted interventions during adolescence can break these cycles by fostering aspirations, enhancing non-cognitive skills, and promoting gender-equitable norms. Life skills programs for adolescent girls, such as those evaluated by Bandiera et al. (2020) and Edmonds et al. (2023), significantly improved educational outcomes, delayed early marriage, and enhanced aspirations.

In many LMICs, including India, parents are the primary decision-makers for their children, especially for girls. Engaging parents is, therefore, essential to ensure sustained progress in empowering girls and achieving gender equity. Research indicates that progressive parental

attitudes can transform outcomes for adolescent girls. Dhar et al. (2019) show that progressive gender attitudes among parents result in higher aspirations for their daughters and greater comfort in gender-equal interactions. Similarly, Giacobino et al. (2024) highlights how financial support for girls' education in Niger, coupled with parental engagement, reduced school dropout and early marriage rates. By targeting both adolescent girls and their parents, our study aims to create a supportive environment that amplifies the impact of the interventions.

The setting of residential schools makes this study pertinent beyond India. Across South Asia and other LMICs, large numbers of adolescents live and study in residential schools, often in environments where educational quality and socio-emotional support are uneven (Porecha, 2024; Asadullah et al., 2009). Many charitable and faith-based organizations run hostels or boarding schools for children across the region. China has an extensive network of residential schools serving rural and semi-urban populations. The scale is substantial: India's KGBV hostel system currently enrolls more than 750,000 disadvantaged girls, while more than 30 million compulsory-school-age students attended boarding schools in China in 2016 (Song, 2020). Despite their scale, these systems frequently struggle to provide proper academic support (Wang et al., 2016; Luo et al., 2009). By testing a low-cost, scalable program within government-run girls' hostels, this study generates evidence directly relevant for policymakers seeking to improve cognitive and non-cognitive skills and gender-related outcomes for adolescents living away from home in resource-constrained boarding school settings.

We find three main results. *First*, the intervention improves several dimensions of girls' human capital. Pooling the two treatment arms, the program significantly increases English proficiency among students in Grades 9–10 by 0.16 SD, confidence by 0.13 SD, and marginally increases aspirations by about 0.09 SD. The cognitive effects are larger under the combined student-and-parent intervention. T2 increases English test scores among Grades 9–10 students by 0.20 SD and final examination marks by about 0.12 SD. The point estimate for aspirations is also larger under the T2 intervention package, at 0.17 SD. However, the estimated T1 and T2 effects on the primary outcomes are not statistically different. We find little evidence of effects on gender attitudes and socio-emotional learning.

*Second*, the intermediate outcomes indicate that the program altered the ways girls invested in their education. The combined intervention led to a substantial increase in educational engagement. T2 students were 7.8 percentage points more likely to report no school absence and spent approximately 1.2 additional hours per week studying. They also reported fewer gender- and family-related barriers to continuing their education. The reduction in perceived family barriers is significantly larger under T2 than under T1. Treated students reported improved peer relationships, increased ability to approach wardens for academic assistance,

greater openness in discussing traditionally sensitive issues, and higher marriage agency.

*Third*, the intervention shifts parental beliefs about the importance of girls' education and confidence that their daughters would achieve their educational and professional goals. Assignment to either treatment arm increases the perceived importance of girls' education by 0.36 SD, confidence in their daughters' ability to achieve educational goals by 0.12 SD, and to achieve professional goals by 0.11 SD. Parents also anticipate higher earnings associated with increased levels of girls' schooling. Many of the parental effects are observed in T1, although T1 parents received no direct intervention. This hints that information or beliefs may be transmitted from daughters to their families. This pattern is consistent with knowledge and behavior being transmitted from daughters to their families (Bhattacharya et al., 2026; Gunawardena et al., 2016; Alix-García and Knittel, 2025).

The findings also show that substantial human-capital gains can be generated cost-effectively under multidimensional household and institutional constraints. The per-student cost for the combined intervention is approximately US$13. A 0.20 SD effect on English test score corresponds to 0.25 learning-adjusted year of schooling (LAYS), equivalent to one quarter of a year of high-quality learning.

This study contributes to the literature on human-capital formation among adolescents in three ways. First, it provides experimental evidence that human-capital gains can be achieved even when children face chronic constraints at both household and institutional levels. The girls in our study come from historically disadvantaged communities living in government hostels with limited academic support. The findings show that even a modest investment under enduring constraints can facilitate human-capital formation without directly addressing these constraints, which are difficult to change in the short run.

Second, the study focuses on both cognitive and non-cognitive domains of human capital. The experimental literature on adolescents to which our study is closely related has focused on either one (e.g., Dhar et al., 2022; Edmonds et al., 2023; Giacobino et al., 2024; Ferrara et al., 2012; Andrew et al., 2022). The study design allows us to examine whether improvements in academic skills are accompanied by changes in the beliefs and behaviors that shape educational investment.

Third, the study produces evidence from an overlooked institutional setting. Existing experimental work on adolescent girls provides evidence from traditional schools, communities, and clubs (e.g., Bandiera et al., 2020; Dhar et al., 2022; Edmonds et al., 2023; Giacobino et al., 2024). There is little evidence from government residential institutions, which suffer from multidimensional constraints. KGBV hostels relax some economic constraints by providing accommodation, food, and education while potentially creating others, such as separation from families and dependence on suboptimal educational and socio-emotional inputs. Our

results show that an intervention delivered under these existing constraints can contribute to human capital formation without employing additional personnel. This makes the study relevant to under-served boarding-school systems across the world.

The effects also extend beyond institutions. Parents in the student-only treatment arm received no direct intervention, yet they exhibit remarkable changes in several outcomes. The result suggests that programs aimed at changing children's human capital may affect household beliefs even without directly intervening with parents.

The rest of the paper is structured as follows. Section 2 outlines the study context, intervention, experimental design, and conceptual framework. Section 3 presents the data, outcomes, and empirical strategy. Section 4 addresses program implementation and fidelity. In Sections 5 through 8, we present the main results, examine potential mechanisms, analyze heterogeneity and program exposure, and report robustness checks. Section 10 considers policy implications and cost-effectiveness, while Section 11 summarizes the findings and concludes.

# 2 Context, Intervention, Design and Conceptual Framework

In India, underprivileged communities such as the Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Classes (OBC) face systemic exclusion that impacts various aspects of their lives, including education.[3] Several studies provide a comprehensive understanding of the challenges and progress in educational attainment among these disadvantaged populations (e.g., Choudhary and Singh, 2017; Dunn, 1993; Goel and Husain, 2018; Mitra, 2008). Literacy rates among these groups are largely low (Raghavendra, 2020; Saravanakumar and Palanisamy, 2013). For men and women not belonging to SC, ST, or OBC categories, literacy rates stand at 91% and 81%, respectively. In contrast, literacy rates are 80% for men and 64% for women among SCs, 78% for men and 61% for women among STs, and 84% for men and 69% for women among Other OBCs (Thakur, 2020). Girls from these communities face a dual disadvantage due to the intersection of caste and gender, resulting in lower educational attainment compared to both SC/ST/OBC men and non-SC/ST/OBC individuals (Rao, 2014). These girls are typically less likely to enroll in traditional schools, experience higher dropout rates, and receive a lesser quality of education than their peers

[3]Scheduled Castes (SCs), Scheduled Tribes (STs), and Other Backward Classes (OBCs) are groups in India identified by the Constitution for special support because they have historically faced social and economic discrimination. SCs and STs are among the most marginalized, affected by caste discrimination to a level of untouchability. According to the Indian last census (Census India, 2011), ST and SC constitute 17% and 8% of India's population, respectively. While the census does not enumerate the OBC population, estimates suggest it accounts for around 35% (Starr and Sahgal, 2021).

from higher socioeconomic backgrounds (Desai and Kulkarni, 2008).

The challenges females face in India extend beyond education, impacting various aspects of life. Early marriage is common, and many women have limited agency (Kishor and Gupta, 2004). Early marriage is more prevalent in disadvantaged communities, limiting educational attainment and career opportunities for females within these communities (Field and Ambrus, 2008; Sunder, 2015). Studies indicate that India's female labor force participation rate (FLPR) remains below the desired level (Afridi et al., 2018; Field et al., 2010, 2021; Klasen and Pieters, 2015). India's current FLPR is 35%, much lower than China's 60% (ILO, 2024). Discriminatory practices within households exacerbate gender inequalities (Ghatak et al., 2024). Cultural expectations often confine women to domestic roles, limiting their aspirations and mobility. Moreover, parental biases and societal expectations undervalue girls' education, perceiving it as a low-return investment (Dhar et al., 2019; Maertens, 2013). Caste-based discrimination further worsens these disparities, deepening economic and social exclusion (Deininger et al., 2013). These constraints restrict education and employment opportunities for girls from backward populations, perpetuating intergenerational inequalities.

The Government of India established the Kasturba Gandhi Balika Vidyalaya (KGBV) program, providing residential schooling to adolescent girls from SC, ST, and OBC communities, as well as orphans and those from below-the-poverty-line (BPL) families to address the educational needs of adolescent girls from disadvantaged communities in 2004. Approximately 5,000 government-funded KGBV hostels operate in educationally backward blocks across the country. The setting for this study is Karnataka, a southern state in India, where 218 KGBV hostels operate with funding from both central and state governments. In Karnataka, SC, ST, and OBC groups collectively account for nearly 80% of the population (Census India, 2011; DoSJE, 2024). KGBV hostels operate under different institutional models designed to address educational barriers faced by girls at different stages of schooling. Type 1 institutions function as residential schools where girls both study and reside on campus, while Type 4 and Type 5 institutions provide residential hostel facilities for girls attending nearby government schools. Type 4 hostels primarily cater to girls enrolled in Grades 9–12, a critical transition period during which girls face an increased risk of school dropout and early marriage. By providing residential accommodation close to secondary schools, these hostels seek to improve retention and facilitate continued participation in higher grades. Type 5 hostels similarly provide residential facilities for girls attending upper-primary and secondary schools in nearby areas. Across all KGBV hostel types, residential schooling seeks to reduce barriers arising from distance, transportation, safety concerns, and limited educational opportunities in remote and educationally backward blocks.

Unlike conventional day schools, girls enrolled in KGBVs reside within the hostel throughout the academic term, returning home only during scheduled holidays. Consequently, the hostel

becomes the primary environment in which students live, study, and have their social lives. A typical day begins with morning prayers, meditation, or supervised study sessions before students attend school during the day. Our field visits and discussions with hostel wardens revealed that evenings were primarily characterised by supervised homework and study sessions, which constituted the principal organised activity outside regular school hours. Beyond these sessions, extracurricular and recreational activities were generally informal and varied considerably across hostels depending on the initiative and engagement of individual wardens. As a result, students' experiences outside formal classroom instruction differed substantially across institutions.

Within this residential environment, hostel wardens occupy a central position in students' daily lives. In addition to managing hostel administration, wardens supervise study hours, monitor students' health and well-being, provide emotional support, and communicate regularly with parents. Particularly for younger girls living away from home, wardens often function as caregivers and trusted mentors, becoming the adults with whom students interact most frequently during the academic year.

Although girls spend most of the academic year in hostels, parents continue to influence important educational and life decisions. Many families belong to rural households engaged in agriculture or daily wage labour and often reside considerable distances from the hostels, limiting opportunities for regular face-to-face interaction. Communication generally occurs through periodic phone calls, visits during holidays, or through hostel wardens. Nevertheless, parental involvement remains important. Baseline survey data indicate that 56.05 percent of students reported helping their families with work even during the academic semester, suggesting that girls continue to balance educational responsibilities with household obligations despite residential schooling.

The institutional characteristics of the KGBV system provided three important considerations that shaped the design of the intervention. First, because KGBV hostels predominantly enroll girls from socio-economically disadvantaged backgrounds, the residential setting provided an opportunity to deliver structured after-school English and life-skills programs to students who were likely to benefit the most from additional academic and socio-emotional support. At the same time, the hostel environment reduced several immediate barriers to school participation, including transportation constraints, safety concerns associated with traveling long distances, and limited access to educational resources.

Second, the central role of hostel wardens in students' daily lives made them natural facilitators of the intervention. Their sustained interaction with girls and established relationships within the hostel ecosystem positioned them to deliver program activities that extended beyond conventional classroom instruction and reinforced learning through regular contact.

Third, while parents interacted less frequently with daughters during the academic year because of the residential nature of the programme, they continued to influence educational aspirations and major life decisions. The baseline evidence on students' continued involvement in household work further highlighted the importance of parental engagement. This motivated the use of an Interactive Voice Response (IVR) platform to strengthen parent-daughter educational support despite limited opportunities for regular in-person interaction.

These contextual features of the KGBV ecosystem and student-parent-warden dynamics informed the design of the intervention, which is described in the following section.

## 2.1 Intervention Details

We partner with *SwaTaleem Foundation*, a non-profit organization, working with adolescent girls living in KGBV hostels, their parents, and hostel wardens with a focus on critical life skills and foundational English lessons. The entire program is approved by the state-level educational authorities. Different components of interventions are detailed below.

The wardens of the KGBV hostels are trained and then serve as teachers in implementing English and critical skills interventions. They provide both academic support and socio-emotional guidance. The program included a 4-day residential training before the intervention started. The *SwaTaleem* Foundation's trainers led the sessions. The training focused on effective communication techniques and interactive lesson delivery methods to improve students' engagement. The training covered aspects of student motivation, pedagogy, and classroom management. A second training was held 6 months after the intervention began, and a third was held 2 months before the intervention ended. In addition to these three formal trainings, the wardens also receive weekly coaching calls from the program staff to review progress and provide feedback. The wardens follow an engagement planner that outlines the structured weekly implementation schedule. The program also includes a non-monetary incentive to motivate wardens.

### 2.1.1 Student-level

***Foundational English lesson***: The English component is targeted towards providing foundational English lessons to girls. The lessons follow a structured, activity-based approach to improve English proficiency, covering basic vocabulary, sentence formation, and reading comprehension. The wardens conduct two 40-minute English sessions per week.

***Life-skills lessons***: The Life-skills component addressed gender sensitivity and breaking stereotypes by distinguishing sex and gender, challenging societal norms, and encouraging girls to aspire beyond traditional roles. The girls were taught to question restrictive gender expectations. They received education on menstrual health, personal hygiene, and balanced

nutrition. It also focused on decision-making and goal setting. The wardens conducted one 40-minute life-skills session per week.

The supplementary Appendix outlines the key components of the Foundational English (Table S1) and Life-skills lessons (Table S2) modules.

### 2.1.2 Parent-level

This component of the intervention aimed to educate, engage, and influence parents by delivering weekly audio messages on key topics via IVR. The messages emphasized girls' education, gender norms, marriage, mental health, and opportunities available to girls. The messages used vignette-style stories that provide both information and encouragement. The consolidated weekly informational contents of the IVR callS are given in Table S3. Each IVR audio call lasted approximately two to three minutes. An automated IVR system called parents every Sunday evening between 7:00 and 8:00 PM to deliver the audio message. If the call went unanswered, the system retried up to three times. If the parents did not answer, a follow-up call is made the following day at 11:00 AM. A call was considered successful if it was heard at least 50% of the full length. If this threshold was not met, an additional scheduled call was made at the end of the week. We tracked the success rate of each week's calls.

## 2.2 Conceptual Framework

The theory of change primarily relies on the hypothesis that improving girls' beliefs, shifting parental attitudes and practices, and enhancing wardens' capacities will jointly raise girls' cognitive skills, confidence, aspirations, and gendered norms. Over the evaluation period, we expect effects on these outcomes, which are determinants of investment in education.

In the short run, exposure to English activities is expected to raise English proficiency. Engagement with life-skills activities is expected to improve students' self-awareness, planning ability, socio-emotional competencies, and sense of agency. As students experience small successes in lessons, their perceived ability, growth mindset, locus of control, and self-esteem should improve, and anxiety may decline. These changes in skills and self-beliefs are expected, in turn, to raise students' aspirations and confidence regarding education, work, and marriage. The more interactive role of wardens is also expected to improve the perceived quality of the student–warden relationship and to create a more supportive hostel environment.

The parent intervention operates through three main channels. *First*, it provides information that can correct misconceptions about the returns to schooling, the feasibility of continued education for girls, and the opportunities available to educated women. Stories of women

from similar backgrounds, coupled with information on scholarships, government schemes, and skills programs, are intended to update parents' beliefs about the returns to girls' education and to lower perceived financial and social constraints. Parents who update these beliefs are expected to be more willing to invest in schooling and open to a formal job outside the home in the future. *Second*, the messages attempt to shift perceived social norms by describing communities and parents who support girls' education, later marriage, and women's work. If parents come to believe that such behavior is more common than previously thought, the perceived social cost of keeping daughters in school and delaying marriage should fall. *Third*, the messages encourage supportive behaviors, such as encouraging study, discussing daughters' goals, and maintaining regular communication with daughters and wardens. These channels correspond to the main parent-level constructs we measure: beliefs about the returns to girls' education, gendered norms, importance of education, and self-reported supportive behaviors.

Wardens receive training in delivering lessons on foundational English and critical skills, along with structured materials and follow-up support. As wardens serve as the primary implementers of the student intervention, these inputs are expected to increase the frequency, regularity, and quality of lessons delivered. Formal and informal training may alter wardens' own beliefs, self-perceptions, and practices. As they gain tools for lesson delivery and observe improvements in students' engagement and performance, their self-esteem and perceptions of students' capability may improve. These changes in wardens' beliefs, stress, and behavioral practices are expected to make the hostel environment more conducive to learning.

Figure 1 below illustrates the impact pathways of the intervention, from inputs to impacts, with the arrows depicting the causal links.

## 2.3 Experimental design

This study employs a cluster randomized controlled trial design. Of the 218 KGBVs in the state of Karnataka, we select 205 hostels before randomization to form our study sample.[4] These 205 hostels are then randomly assigned to three groups within each strata: two treatments (T1 and T2) and a control (C). The strata are defined as a combination of administrative regions and type of KGBV schools, leading to 12 strata.[5]

***Treatment Arm 1 (T1: Learning Program).*** Students assigned to this treatment arm received structured English and life-skills lessons delivered by hostel wardens.

[4]We excluded 4 where the pilot survey was conducted, and excluded 9 that had fewer than 25 students.

[5]KGBV Type 1 are hostels where schools and hostels are within the same premises, and have grades 6 to 8. KGBV Type 4 are hostels with grades 9 to 12, and KGBV Type 5 are with grades 6 to 10. Both KGBV Type 4 and 5 are independent hostels, and students are enrolled in the nearby schools.

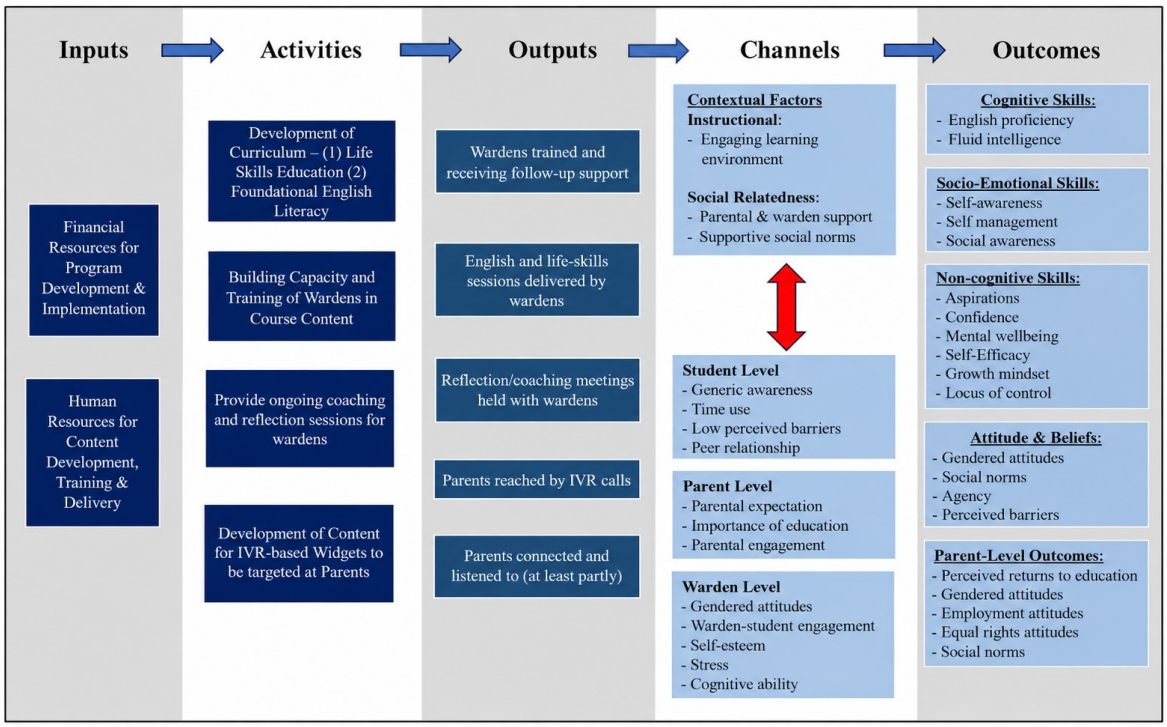



Figure 1: Theory of change: Inputs to outcomes for adolescent girls and parents

***Treatment Arm 2 (T2: Learning Program + Parental IVR).*** Students assigned to this treatment arm received the same English and life-skills lessons as those in T1. In addition, their parents received weekly interactive voice response (IVR) audio messages designed to reinforce the intervention and encourage parental engagement in their daughters' education.

***Control Group (C): Business-as-Usual.*** Students assigned to the control group received no intervention and continued with the standard KGBV curriculum and activities.

Figure 2 shows the geographical distribution of the KGBV hostels randomized in the study.

## 2.4 Timeline

We conducted a pilot survey in four selected KGBVs in mid-August 2024. Wardens received an initial four-day in-person training in mid-October 2024. The interventions started in the second week of November 2024 and ended in the first week of November 2025. The endline survey for students started in the fourth week of November 2025, and the parent endline survey began two weeks after. Figure 3 provides a detailed visual representation of the project timeline.

# 3 Data collection, Outcomes, and Empirical Strategy

We collected data at three key points: baseline, during the intervention, and endline. We collected baseline data before randomization.

## 3.1 Surveys: Baseline and Endline

**Baseline survey** We administered a baseline survey with students, parents, and wardens. The student survey took place in a classroom setting using pen and paper, on OMR (Optical Mark Recognition) sheets. Students completed a multiple-choice survey consisting of foundational English, numeracy, and critical general aptitude; perceived barriers to education; aspiration; growth mindset; and mental well-being. In addition, we gathered demographic information, including household characteristics and educational backgrounds.

**During-intervention data** We collected implementation data to assess the intensity and compliance in hostels. We conducted two in-person classroom observations using a standardized tool to capture the quality of lesson delivery. For the IVR intervention, we collected data on how many parents listened to the messages and the extent of their engagement.

**Endline survey** At the endline, we surveyed students, parents, wardens, and the siblings of program students. These surveys included items from the baseline and additional measures. We also administered three incentivized tasks with students to measure their private, public, and leadership confidence.

## 3.2 Sample

The student population comprises all students enrolled in Grades 6–9 in the academic year 2024-25 in the KGBVs randomized into the trial. We randomized 205 KGBVs. However, two of them could not be surveyed because they were not operational. At baseline, we surveyed 13,383 students from 203 KGBVs. During the intervention, one control KGBV merged with a nearby treatment KGBV. So we have 202 KGBVs at endline. We surveyed 15,360 students at endline, including the new admits in Grades 6–10. Students enrolled at baseline had advanced by one grade. Thus, the target study cohort was in Grades 7–10 at endline. Of the endline students, 9,243 could be matched to the baseline survey. Our main ITT analysis is based on this matched sample.

For the parent survey, we randomly (on-the-spot lottery) selected 4,802 parents across treatment and control groups at baseline, of whom 4,216 were tracked at endline.

We also surveyed wardens from the KGBVs. We could not survey warden from one KGBV as she was absent on the survey day. One KGBV was merged with a nearby KGBV after

baseline. Thus, the warden sample comprises 202 observations in both baseline and endline; 158 of them participated in both baseline and endline surveys. Table A1 summarizes the baseline, endline, and matched samples for students, parents, and wardens.

Table A2 and Table A3 presents baseline balance across student, household, warden, and KGBV characteristics. Overall, the treatment and control groups are well balanced. Joint F-tests fail to reject the null hypothesis of equality of baseline characteristics across the treatment and control groups for student, household, and warden characteristics. At the KGBV level, we find a statistically significant joint difference between T2 and the control group. However, most individual differences are small in magnitude. At the student level, the treatment groups are comparable across demographic characteristics, grade composition, attendance, study time, cognitive ability, aspirations, perceived barriers, and socio-emotional measures. We observe a few statistically significant differences in baseline academic outcomes, particularly in English test scores among T1 students and in previous examination marks among T2 students. Our main econometric specifications control for baseline student characteristics, including prior academic achievement.

## 3.3 Output and Outcome Variables

Our output and outcome measures follow the contextual framework outlined earlier. Under the outputs, we have measures on total training sessions attended by wardens, number of lessons organized by wardens in their hostels (reported by students), delivery of structured lessons by wardens, quality of teaching practices from our classroom observations, reach and intensity of engagement of calls made to parents from the IVRS dialer dataset.

We pre-specified primary outcomes at two levels: student and parent. We construct the indices using Principal Component Analysis (PCA) and then standardize with a control mean and standard deviation. All outcome variables, their response scales, and index constructions are provided in detail in the Supplementary Appendix.

*Cognitive outcomes*: Our first outcome is the English test score, which is closely related to the intervention. We administered two MCQ-type English tests: one for students in Grades 6–8 and a more difficult version for students in Grades 9–10. Each test had seven questions, each worth one point. English test score is the total point achieved in the test. We also asked the Raven's test to assess fluid intelligence. Further, the survey asked students to report their school-level academic test score from the previous year.

*Non-Cognitive outcomes*: Our first measure is *confidence index* which is a simple mean of three components: leadership confidence, public confidence, and private confidence. Each component is elicited through incentivized behavioral tasks. Leadership confidence captures students' intention to become class monitor. Before making the decision, they were informed

that one student would be randomly selected to deliver a one-minute speech on his/her candidacy. The selected student receives three gift tokens for delivering the speech adequately, one token for making only a limited attempt, and no token for refusal or being unable to speak. Through similar incentives, public and private confidence is elicited. Public confidence captures whether a student opts to solve a puzzle in front of the class, while private confidence is a measure of their prediction score on the cognitive tests. We construct an *aspirations index* covering three domains following Beaman et al. (2012): (i) educational (dummy=1 for students aspiring to study beyond school) , (ii) occupational (dummy=1 if students opt to become doctor, engineer, government officer, police or defense officer, or politician), and (iii) desired age to marry is above age twenty-five. Other non-cognitive outcomes are: *locus of control* following(Alan and Kubilay, 2025), *growth mindset* following(Sisk et al., 2018), *marriage agency index* (with questions like "girls should choose on their own when to marry"), openness to respond to menstruation-related questions, and sharing of emotions.

Within the broad non-cognitive domains, we also create *socio-emotional skills index* using twelve items adapted from Anthony et al. (2020) and Ganicho et al. (2025). The items cover three domains each on a four-point Likert scale: self-awareness, self-management (staying calm while dealing with problems, disagreeing with others, and paying attention when the teacher talks), and social awareness (understanding others' feelings, standing up for others, helping friends in need, etc.). We also create *social anxiety index* outcome from 5 items measuring how much fear/avoidance students feel in different situations such as attending social gatherings, speaking up in classroom, presenting in classroom, speaking with authority figures etc.). We also create outcomes for overall well-being measures - getting headaches/sickness feeling, trouble concentrating in studies, or feeling of general anxiety.

*Attitudes and Beliefs*: We construct *gender attitude index* using a 9-item inventory, following Dhar et al. (2022), to measure students' gender attitudes. These statements cover attitudes toward women's roles within the household, educational opportunities for girls, equal rights, political leadership, and gender-based violence. Survey items are recoded where necessary so that higher index values indicate more progressive gender attitudes. Similarly, we create the *social norms index* capturing students' perceptions of prevailing community norms regarding women's education, employment, and gender roles.

*Other student-level measures*: We create several important intermediate outcomes (or potential channels) following our contextual framework. First, broad category is students' support network, of two separate relationship-based indices measured on five-point Likert scale: *peer relationship* (5 questions, e.g- "I feel safe in the hostel", "I like to study with other hostel mates") and *student-warden relationship* (5 questions, e.g- "I enjoy spending time with warden", "I like my warden"). We also explicitly measure whether students feel comfortable

approaching the warden for any study-related help. To measure relationship with parents, our outcome is whether the student calls family at least once a week, and how comfortable they feel sharing hostel/school experiences on a scale of 1 to 5 (1- not at all comfortable to 5- very comfortable). Second, the broad category of measures is inspired by the contextual model of engagement Lam et al. (2012) and used in Bharti et al. (2026). We capture proximate factors known to influence students' academic engagement, such as *perceived barriers*, behavioral engagement (school absenteeism), and cognitive engagement (studying hours on weekdays and weekends).

*Parents'-level outcomes*: Since parents are a crucial pillar of our intervention, we also create parents'-level outcomes. Under attitudes and beliefs of parents, we create gender attitude index and social norms index outcomes, following the same inventory described above. We also create two other gender-related indices, asked on a 5-point Likert scale: *employment attitudes* consisting of three items (such as "men are better suited than women to work") and *Equal rights attitude* with two items ("parents should maintain stricter control over their daughters than son", and "girls should attain higher education so that they find better students"). We construct a standardized *importance-of-education index* spanning four questions about how a college degree would impact their daughter's success, job prospects, social status, and financial independence. Each item is asked on a scale of 1 to 5 (1- not at all likely, and 5- very likely). We also create outcomes measuring parents' confidence in their daughters' educational and professional achievement. Our other important measure is parents' expected earnings for their daughters after the completion of Grade 8, Grade 10, Grade 12, or college.

Our primary outcome variables (as in the pre-analysis plan *AEARCTR-0015254*) at the students' level are: English test score, confidence index, aspiration index, gender attitudes index, and socio-emotional skills index. At the parents' level, our primary outcomes were the importance of education index and confidence in their daughters (expected earnings). The rest are either secondary or exploratory variables.

## 3.4 Estimation strategy

We estimate intention-to-treat (ITT) effects using ordinary least squares (OLS). Our analysis follows two steps. First, we estimate the average effect of assignment to any treatment arm relative to the control group. Second, we estimate the effects of *T1(Learning Program)* and *T2(Learning Program + Parental IVR)* separately. Unless otherwise stated, all regressions include pre-treatment student-, household-, warden-, and KGBV-level controls, as well as randomization-stratum fixed effects. Standard errors are clustered at the KGBV level.

Estimating the effect of pooled treatment increases statistical precision when the two inter-

ventions have similar impacts, while estimating arm-specific effects allows us to assess the incremental contribution of parental engagement.

For the pooled treatment effect, we estimate

$$Y_{ih,t=1} = \alpha + \beta T_h + \theta X'_{ih} + \phi H'_h + \lambda W'_h + \delta s + \gamma Y_{ih,t=0} + \varepsilon_{ih} \tag{1}$$

where $Y_{ih,t=1}$ denotes the endline outcome for individual $i$ (student, parent, warden) associated with KGBV $h$. $T_h$ is an indicator for assignemnet of KGBV $h$ to any of treatment arms. The coefficient $\beta$ captures the average ITT effect of assignment to any of the interventions.

To estimate arm-specific treatment effects, we utilize

$$Y_{ih,t=1} = \alpha + \beta_1 T_{1h} + \beta_2 T_{2h} + \theta X'_{ih} + \phi H'_h + \lambda W'_h + \delta s + \gamma Y_{ih,t=0} + \varepsilon_{ih} \tag{2}$$

where $T_{1h}$ and $T_{2h}$ are indicators for assignment of KGBV $h$ to Treatment 1 and Treatment 2, respectively. $Y_{ih,t=1}$ is outcome at level $i$ (student, parent or warden) associated with KGBV $h$ in the endline survey. The coefficients $\beta_1$ and $\beta_2$ provide ITT estimates of the impact of Treatment 1 and Treatment 2, respectively, relative to control.

The vector $X'_{ih}$ contains pre-treatment student and household characteristics, $H'_h$ contains baseline KGBV-level characteristics, $W'_h$ contains baseline warden characteristics, and $\delta s$ are strata fixed effects. $\varepsilon_{is}$ is the error term. Where a comparable baseline measure of the outcome is available, we include it as $Y_{ih,t=0}$.

To examine differences between the treatment arms (T1 and T2), we report the *p-value* of the null hypothesis $\beta_1 = \beta_2$.

# 4 Implementation Output

This section summarizes implementation fidelity with respect to the key outputs specified in the program's theory of change: preparation and support of wardens, delivery of the student intervention, and delivery and receipt of the parent IVR intervention.

**Student Intervention.** Wardens were the primary facilitators of the student intervention, making their training and continued coaching an important first stage of implementation. Across the 133 treatment hostels, wardens participated in an average of 3.32 of four days in Workshop 1, 3.61 of four days in Workshop 2, and 2.07 of three days in Workshop 3. Wardens from treatment hostels also participated in an average of 17.91 of the 25 coaching

calls conducted over the two intervention cycles. Attendance is measured at the KGBV level and captures participation by any warden from the KGBV during the intervention period (Appendix Table A4).

Student reports indicate substantial, although incomplete, exposure to the curriculum. Students attended an average of 31 English and 15 life-skills sessions, corresponding to approximately two-thirds of the planned sessions. Between 41 and 54 percent of students reported that each life-skills module had been taught, with a similar proportion reporting active participation. Among students who received the modules, comprehension was high, with approximately 90% reporting that they understood the lesson content across all five domains (Appendix Table A5). Consistent with these reports, treatment wardens were 25 to 35 percentage points more likely than control wardens to report delivering structured lessons across several non-cognitive domains, including self-awareness, social awareness, life skills, health and hygiene, and gender sensitivity (Table 1).

Independent classroom observations provide further evidence on the quantity and quality of delivery. Sessions were successfully observed in 124 KGBVs in Round 1 and 98 in Round 2; the smaller Round 2 sample primarily reflects a higher incidence of scheduled sessions being canceled or not taking place on the observation day. Wardens were on task for an average of 83.74 percent of observed class time in Round 1 and 77.38 percent in Round 2, while the corresponding student time-on-task rates were 78.90 and 74.49 percent. Mean scores across observed teaching-quality dimensions ranged from 5.06to 6.88 in Round 1 and 5.53 to 6.91 in Round 2 on a 0–10 scale (Appendix Table A6). Taken together, these measures indicate substantial implementation of the student intervention, while also documenting variation in programme exposure and delivery across hostels. This provides an important first stage for interpreting the student outcomes examined in the subsequent results.

**Parent Intervention.** The parent intervention was delivered through weekly Kishori Vani IVR calls over 52 weeks. We distinguish between *reach*, defined as answering a call (recorded call duration greater than zero), and *successful engagement*, defined as listening to at least 50 percent of the audio message. Across the 52 intervention weeks, an average of 74.37 percent of targeted parent phone numbers answered the IVR calls, while 32.58 percent listened to at least half of the audio message. Appendix Table A7 reports the weekly implementation statistics, while Appendix Figure 6 shows how these engagement measures evolved over the intervention period.

Parent endline reports and administrative call records provide complementary measures of individual-level exposure, which are given in Table A8. Among 1,384 surveyed parents in the parent-treatment arm, 56.4 percent reported that they or a relative had received a Kishori Vani call, while 28.7 percent reported having the Kishori Vani number saved. Overall, 57.4

percent reported at least one of these two indicators of exposure. Of the 1,384 surveyed parents, 1,163 (84.0 percent) could be matched to the administrative IVR records. Among these matched parents, exposure varied substantially: 8.3 percent never listened to at least half of an audio message, while, on average, parents reached the 50-percent listening threshold in approximately 15 intervention weeks. Thus, the IVR intervention achieved broad reach, but the intensity and persistence of engagement varied considerably across parents.

# 5 Results

This section presents the ITT effects of the two interventions on students' cognitive and non-cognitive outcomes. We report three sets of estimates. The "any treatment" coefficient compares students assigned to either intervention arm with students assigned to the control group. The T1 (Learning Program) and T2 (Learning + Parental IVR) coefficients show the experimental distinction between the lesson-only and lesson-plus-IVR treatments. We also report the p-value from a test of equality between the two treatment-arm coefficients. Table 2 reports the primary student outcomes, pre-specified as primary outcome domains. Additionally, we report final marks from centrally administered standardized school examinations, pre-specified as secondary outcomes, as they are informative about whether any program effects extend beyond the study-administered English test. [6]

## 5.1 Cognitive and academic outcomes

The first three columns of Table 2 report effects on the English assessments and final examination marks. Students in the younger and older grade groups were provided with different English tests with different difficulty levels. We therefore present the English effects separately for Grades 9–10 (Senior group) and Grades 7–8 (Junior group). This separation follows the ex-ante measurement design rather than an ex-post division of a common assessment.

**English Test Score.** The findings for students in Grades 9–10 are positive and statistically significant at the 5% level. Pooling the two treatment arms, assignment to the program increases English performance by 0.161 SD, with a standard error of 0.079 and a p-value of 0.044. The evidence shows a positive effect of the complete T2 package among older students. The arm-specific estimates are positive for both interventions. T1 increases the English score by 0.124 SD, although this estimate is imprecise. T2 increases the score by 0.200 SD, with a standard error of 0.091 and a p-value of 0.029. Thus, we find evidence that the intervention improved performance on the English test administered to the older grade

[6] The results reported use students' self-reported final examination marks from the 2024–25 academic year; the exams were conducted approximately four months after the intervention began. We plan to collect their actual marks from the government authority and update the analysis using the 2025–26 final examination marks once these data become available.

group. However, the T1 and T2 estimates are not statistically different.

For students in Grades 7–8, we fail to reject the null hypothesis that the English test score is not different from zero. The estimated pooled treatment effect on English performance is 0.020 SD (p=0.869). The estimate is close to zero and statistically insignificant. The arm-specific coefficients are also small relative to their precision: 0.061 SD for T1 and -0.022 SD for T2. The T1 and T2 estimates are not statistically different. The confidence intervals for this outcome are also wide.

**School Final Marks.** Final examination marks provide complementary evidence of an academic effect. The pooled treatment estimate is positive, at 0.07 SD, but statistically insignificant. The T1 effect is close to zero, at 0.018 SD. In contrast, T2 increases final marks by 0.123 SD, with a p-value of 0.042. The p-value for equality between T1 and T2 is 0.079, providing suggestive evidence that the combined intervention was more effective than lessons alone for this outcome. Final marks are meaningful because they are not based on a test developed specifically for the intervention. An effect on centrally administered school examinations indicates the T2 benefit may have extended beyond familiarity with the English curriculum used by the program. It also confirms that the program did not crowd out students' performance on school-based assessments; rather, it improved it.

The broader cognitive table in Appendix Table A9 also shows the pooled English score and Raven matrices performance. The pooled English estimate is positive but statistically insignificant. This is due to a positive senior-grade effect combined with an effect close to zero in the younger-grade group. Neither intervention produces a precise estimate for the student's fluid intelligence. This pattern suggests that the academic improvements are more closely related to acquired skills, examination preparation, or study behavior than to a general change in nonverbal cognitive ability.

## 5.2 Non-cognitive outcomes

The remaining columns of Table 2 report the non-cognitive outcome indices. The results indicate improvements in confidence and aspirations, but limited effects on gender attitudes and socio-emotional learning (SEL).

**Confidence.** Unlike the other non-cognitive outcomes, which are primarily based on self-reported survey responses, the confidence index captures students' behavioral expressions of confidence through three incentivized tasks. The incentives were designed to make students' choices consequential and thereby reduce the scope for socially desirable responses.

The program improves students' confidence index. Pooling both treatment arms, assignment to the intervention increases confidence by 0.130 SD, with a p-value of 0.033. T1 increases

the index by 0.146 SD, significant at the 5% level. The T2 estimate is similar in magnitude but less precise. The two treatment effects are not statistically different (p = 0.705). Both arms generate positive and quantitatively similar point estimates. We, therefore, cannot characterize confidence as a T1-only effect. Splitting the index by different confidence domains, Appendix Table A10 shows that the confidence index is primarily driven by leadership confidence. The pooled treatment effect on leadership confidence is 4.3 pp and is marginally significant. T2 produces a larger leadership-confidence estimate of 5.7 pp (p = 0.057), while T1 produces an estimate of 2.9 pp. However, the treatment-arm difference is not statistically significant. Neither intervention has a precise estimate on the other two confidence measures: public confidence (willingness to solve a puzzle publicly) and private confidence (self-prediction of performance in survey-based tests).

**Aspirations.** The program improves students' aspirations in general and their aspirations to study beyond school-level in particular. Pooling the two treatment arms, the intervention increases the aspirations index by 0.085 SD. The estimate is marginally significant, with a p-value of 0.071. The T1 estimate is positive but statistically insignificant. T2 increases the aspirations index by 0.119 SD, with a standard error of 0.058 and a p-value of 0.043. The equality test fails to reject the null hypothesis that the effects across the two treatment arms are the same (p = 0.268).

Appendix Table A11 decomposes the aspirations index into educational, occupational, and marriage-related aspirations. The most consistent effect concerns the desire to continue education beyond Grade 12. The pooled treatment effect is 4.5 pp at the 1% level (control mean of 77%). The effect among T1 students is 3.8 pp, while it is 5.3 pp among T2 students. Both arm-specific estimates are statistically significant. However, the equality test indicates no meaningful difference between them. In contrast, neither intervention significantly changes aspirations for a high-level occupation or preference for marriage after age 25. The aggregate aspirations effect therefore appears to be driven mainly by continued education rather than by a generalized increase in all future ambitions.

**Gender attitudes.** We find limited evidence of an effect on the overall gender-attitudes index. The pooled treatment coefficient is 0.070 SD and is statistically insignificant. T1 produces an estimate close to zero, while the T2 estimate is positive at 0.154 SD but imprecise. The equality test suggests a possible difference between the treatment arms but falls short of conventional significance (p = 0.091). Decomposing the index in the Appendix Table A12, shows that the effects are concentrated in education-related attitudes. The pooled treatment effect on the education-attitude index is 0.105 SD. T1 increases this index by 0.123 SD, while T2 produces an estimate of 0.087 SD. There is weaker evidence that T1 improves equal-rights attitudes, but no clear effects on the employment-attitude index. Across generations, the program seems to be more effective in shifting beliefs closely connected to

education opportunity while having little impact on more deeply rooted gender norms. In line, we also do not find any consistent effects on the social-norms index in Column (6) of the Appendix Table A14.

**Socio-emotional learning (SEL).** The program has no statistically significant effect on the overall SEL index. The pooled treatment estimate is 0.056 SD. The T1 coefficient is 0.004, while the T2 coefficient is 0.109. The component results in Appendix Table A13 are likewise limited. T2 increases the social-awareness index by 0.136 standard deviations, with a p-value of 0.068. There is no clear effect on self-awareness or self-management sub-indices. The life-skills knowledge measures show some evidence that students retained specific curriculum content (Table A15). The pooled treatment effect increases correct responses to the empathy item by 8.2 percentage points, and T2 raises the same outcome by 10.5 percentage points. There is weaker evidence of effects on gender-role knowledge and goal-setting. Most other content-knowledge outcomes are unaffected. The knowledge results therefore provide partial evidence of content acquisition.

Psychological or emotional distress can manifest as somatic symptoms, including headaches and abdominal pain (Ayonrinde et al., 2020; White and Farrell, 2006). Young students may also be unwilling to respond directly to such questions. Hence, we collected measures such as stress, concentration difficulties, and the anxiety index, but did not find any consistent effects on them (Appendix Table A14).

**Student agency.** Table 3 examines openness about menstruation, emotional sharing, preferred marriage age, marriage agency, and locus of control. Pooling the treatment arms, the program increases openness about menstruation by 4.3 pp, the proportion of students feeling comfortable sharing emotions by 4.4 pp, and the marriage-agency index by 0.123 SD. The pooled locus-of-control estimate is also positive, at 0.090 SD, though only marginally significant.

T1 exhibits the largest effect on menstruation openness, increasing it by 5.3 pp. T2 increases emotional sharing by 5.7 pps and the marriage-agency index by 0.157 SD. T2 also raises preferred marriage age by 0.28 years, slightly above three months, although this estimate is marginally significant. The difference in preferred marriage age between T1 and T2 is statistically significant ($p = 0.044$).

These results suggest that the intervention affected specific dimensions of girls' agency rather than a single generalized construct. The lesson component appears to have increased comfort in responding to a sensitive menstruation question. The combined intervention improved emotional expression and perceived agency over marriage-related decisions. Such effects are consistent with a curriculum that encouraged self-expression, future planning, and discussion of gendered constraints.

# 6 Potential channels

We present outcomes that may help explain the primary students' level findings. We club them into two student-level categories: perceived barriers to education, educational engagement, student agency, and support networks. The second is the parents' level.

**Perceived barriers to education.** Table 4 exhibits students' perceptions of gender, family, caste, and economic barriers, as well as educational self-efficacy. The pooled treatment effect reduces perceived gender barriers by 5.8 pp and is marginally significant ($p = 0.053$). T2 reduces gender barriers by 8.0 pp from a control mean of 50%, significant at the 5% level ($p = 0.035$). T1 produces a smaller and statistically insignificant reduction of 3.6 pp. T2 also reduces perceived family barriers by 7.1 pp from a control mean of 44.7% ($p = 0.050$). The T1 estimate is close to zero. The difference between T1 and T2 is statistically significant at the 5% level. The program could plausibly alter students' perceptions of gendered constraints through the life-skills curriculum. As expected, neither treatment has a precisely estimated effect on caste barriers, economic barriers, or educational self-efficacy. The intervention was not designed to relax household financial constraints or perceived caste barriers.

Reductions in gender and family barriers may help explain the increase in educational aspiration and attitudes. Students may be more willing to envision continued education when they perceive fewer social or household obstacles. The family-barrier result is particularly relevant to the parental-IVR component because it is one of the few outcomes for which the T2 coefficient is both substantively larger and statistically distinguishable from the T1 coefficient.

**Educational engagement.**: The treatment has improved engagement: reducing absenteeism in school, increasing study hours. Table 5 shows that T2 substantially increased educational engagement. Pooling both treatment arms, assignment to the program increases the proportion of students who never miss school by 5 pp (p-value = 0.067). The T1 estimate is 2.4 pp and statistically insignificant. T2 increases the proportion by 7.8 pp ($p = 0.016$), relative to a control mean of 61.6%. This represents an increase of approximately 13% relative to the control mean. The equality-test p-value is 0.068, which provides suggestive evidence that the attendance effect was stronger under T2.

The contrast between the treatment arms is clearer for study time within the hostel. T2 increases weekday (5-day) average study time at the hostel by 0.97 hours and total weekly (7-day) hostel study time by 1.24 hours. Both estimates are statistically significant at the 5% level. T1 shows negative but imprecise estimates for the same outcomes. The T1–T2 differences are statistically significant, with p-values of 0.013 and 0.025. Neither treatment significantly changes weekly study time at home, most likely because students go to their

homes during vacation or festivals.

These results are consistent with the academic findings. The T2 students studied more within the hostel and were less likely to report missing school. At the same time, T2 also generated the strongest effects on senior English performance and final marks. The absence of an effect on study time at home suggests that the engagement response occurred primarily within the institutional environment in which the lessons were delivered. The parental IVR component may have reinforced the importance of attendance and study. The engagement results provide plausible evidence of a mechanism but not direct proof of formal mediation.

**Support networks.** Table 6 reports effects on students' relationships with peers, wardens, and family. Pooling both treatment arms, the program significantly increases the peer-relationship index by 0.135 SD ($p = 0.006$). It raises the proportion of students who can approach the warden for study help by 5.6 percentage points ($p = 0.019$).

Both T1 and T2 significantly improve peer relationships. T1 raises the peer relationship index by 0.127 SD, while T2 raises it by 0.144 SD. The estimates are nearly identical. It suggests that this effect likely arises from the component of common lessons.

T2 produces higher point estimates for warden-related outcomes. Increases the student–warden relationship index by 0.137 SD and increases the proportion of students who consider their wardens approachable for study help by 7.9 percentage points. The relationship index is marginally significant, while the approachability effect is precisely estimated at the 1% level. There is also suggestive evidence that T2 increases students' comfort in sharing hostel experiences. However, direct comparisons between T1 and T2 are not statistically significant.

The results of the support network complement the educational-engagement findings. Stronger peer relationships and greater access to wardens may make it easier for students to seek academic assistance, discuss personal concerns, and remain engaged with academic tasks. The results also indicate that the program operated partly through the institutional and relational environment of the hostel, not solely through transmission of lesson content.

## 6.1 Parent-level

Following our conceptual framework, we explore parent-level outcomes that could potentially help explain student-level impacts. The student- and parent-level results indicate that the program influenced educational investments through complementary changes among both students and their families. Although the two components of the program target distinct groups, the observed behavioral changes are mutually reinforcing and reveal a potential mechanism of impact.

We observe that the parent-level impacts are consistent with students' outcomes. Treated parents place greater importance on girls' education and express more confidence that their daughters would achieve their educational and professional goals.

**Attitudinal Changes.** Column (1) of Table 7 reports the treatment effects on the Importance of Education Index. The intervention substantially improved parents' perceptions of the importance of girls' education. The pooled treatment effect is 0.36 SD, significant at the 1% level. Examining the treatment arms separately, both T1 and T2 generated positive and statistically significant improvements. Although the estimated effect is larger for T1 than for T2, the difference between the two treatment arms is not statistically significant ($p = 0.102$). These findings suggest that improvements in parents' perceptions of the importance of girls' education are largely driven by the student intervention itself, with no additional effect attributable to the parental IVR component. The intervention also improved parents' attitudes toward women's employment. In the pooled specification, assignment to either treatment arm increased the employment attitude index by 0.196 standard deviations relative to the control group. The estimate is statistically significant at the 5 percent level.

However, the intervention did not produce measurable changes in the prevailing gender attitude, social norms or equal rights attitude during the study period. Column (3) in Table 7 reports the effects on parents' gender attitudes. Although the estimated coefficients are uniformly positive, neither the pooled treatment effect nor the treatment-specific estimates are statistically significant. Moreover, there is no evidence that the treatment effects differ across the two treatment arms ($p = 0.524$). Similarly, parents became slightly more likely to believe that girls could perform selected financial and agricultural roles (Table A28) and expressed more support for daughters' employment in certain workplace settings (Table A29). However, these effects are neither broad nor consistent across outcomes.

**Confidence in their Daughters.** The intervention substantially increases parents' confidence in their daughters' educational and professional success (Table 8). Parents in the treatment group are significantly more likely to believe that their daughters would achieve their desired educational attainment and professional goals. Columns (3)–(6) examine parents' perceived financial returns to girls' education. The returns are measured by expected earnings after completing Grade 8, Grade 10, Grade 12, and college. The pooled treatment effects are positive across all educational milestones and are statistically significant for expected earnings after completing Grade 10, Grade 12, and college. These findings indicate that the intervention increased parents' expectations regarding the labor market returns to higher levels of girls' education. When the treatment arms are considered separately, the estimated effects are consistently positive in both groups, although the effects are more precisely estimated under T1. Nevertheless, formal tests fail to reject the equality of the T1 and T2 coefficients for any of the four outcomes. However, the treatment effects on par-

ents' expectations regarding their daughters' timing of key life events such as marriage age, education completion age, etc. are modest. These responses seem to be dominated by the rigid social norms. The expected age to complete education, get married and have a first child in the control group parents are: 22, 23 and 25 years, respectively, and the treatment group parents' responses are no different. All the coefficients are very small (Column 2-4 in Appendix Table A30).

**Parents' Behavioral Engagement.** We examine whether the interventions influenced parental engagement with the hostel. Parents report modest improvements in communication with wardens and daughters (Table A25). Treated parents are more likely to communicate with the hostel warden and their daughters. They are also more likely to report that their daughters felt comfortable sharing hostel experiences. Beyond these improvements, the evidence for broader changes in parental beliefs is mixed. The interventions do not consistently modify parents' expectations regarding their daughters' educational attainment, marriage timing, or future family formation (Table A24). Similarly, parents' knowledge of their daughters' educational aspirations and academic interests changes little. Although parents become somewhat better informed about the hostel environment, such as the hostel location and the warden's name (Table A27). While the treated parents are somewhat more likely to discuss the IVR program with wardens, there is little evidence that the interventions broadly change the topics they discuss with hostel wardens (Table A26).

## 6.2 The Nexus between Student and Parent Outcomes

The student- and parent-level results indicate that the program influenced educational investments through complementary changes among both students and their families. Although the two components of the program target distinct groups, the observed behavioral changes are mutually reinforcing and reveal a potential mechanism of impact.

The student intervention led to improvements in English skills, final examination performance, confidence, and educational aspirations. In addition to these outcomes, there was increased educational engagement, stronger relationships with peers and wardens, and reduced perceived educational barriers. The program also produced modest improvements in broader gender attitudes and socio-emotional learning. These findings indicate that the primary effects are mediated by enhanced academic engagement and students' beliefs about future opportunities, rather than by broad attitudinal change.

The parent-level results exhibit a similar pattern. Parents are more likely to regard girls' education as important, perceive greater economic returns to educating their daughters, express increased confidence in their daughters' educational and professional success, and report substantially higher satisfaction with the hostel environment. The intervention also

modestly strengthened communication among parents, daughters, and wardens. However, there is limited evidence that the program alters parents' gender attitudes or social norms.

Collectively, these findings indicate that the program mainly influenced expectations and educational investments rather than altering deeply rooted social preferences. Students demonstrate increased engagement in learning and greater optimism regarding their future educational trajectories. Simultaneously, parents exhibit increased confidence that investing in their daughters' education would produce meaningful returns. The association between student and parent outcomes is noteworthy, given that parents in T1 did not receive direct intervention. The positive effects observed among these parents suggest that the program's influence may have been transmitted indirectly through students.

In general, the evidence suggests a reinforcing process in which improvement in students' academic engagement and confidence coincides with increased parental support for girls' education. Instead of transforming gendered norms and attitudes within a short timeframe, the intervention appears to have strengthened the educational environment for disadvantaged girls by increasing both students' motivation to invest in their own human capital and parents' willingness to support these investments.

# 7 Heterogeneity and program exposure

We explore whether treatment effects differ by program exposure and parents' education. The dose-response analysis with respect to program exposure is exploratory in nature, rather than causal. Attendance is measured after treatment assignment and may reflect student motivation, health, warden quality, KGBV hostel management, or other characteristics that also affect outcomes. The estimates are therefore most appropriately described as descriptive evidence that treatment effects were concentrated where exposure was greater. We follow this analysis with a formal IV analysis, using assignment to treatment as an instrumental variable.

The attendance results show a strong positive association between the number of English sessions attended and cognitive outcomes (Table A16). The program conducted 46 English lessons and 23 life-skills lessons during the intervention period. Among students attending most English sessions (28–46 sessions), T1 is associated with increases of 0.205 SD in the pooled English score, 0.195 SD in the Grades 9–10 English score, and 0.137 SD in final marks. For high-attendance T2 students, the estimates are 0.207 SD for Pooled English, 0.295 for senior English, 0.190 for final marks, and 0.211 for Raven performance. Similar patterns appear for life-skills attendance. Students in the highest attendance category (14–23 sessions) show larger estimates for confidence, aspirations, gender attitudes, and SEL (Table A17). These results are consistent with stronger implementation generating larger

program effects.

To examine the effects of program exposure, we estimate treatment-on-the-treated (ToT) effects using two-stage least squares. We instrument session attendance with randomized treatment assignment. The first-stage estimates indicate very strong instruments. Assignment to either treatment arm increased attendance by approximately 31 English sessions and 15.5 life-skills sessions.

The IV estimates for cognitive and non-cognitive outcomes are reported in Tables A18 and A19, respectively. The coefficients represent the estimated effect of attending one additional program session for students whose attendance increased because of treatment assignment. Evaluated at the average treatment-induced increase in attendance, the implied effects are similar to the corresponding ITT estimates. Approximately 31 additional English sessions increased the Grades 9–10 English score by about 0.19 SD, while approximately 15.5 additional life-skills sessions increased the confidence and aspirations indices by about 0.12 and 0.11 SD, respectively.

The household-education results suggest that some T2 academic effects may be larger among students from less-educated households (Table A20). We divide household education into two categories: Below Grade 10 if both parents have less than 10 years of education, Grade 10 or above if at least one of the parents has more than 10 years of education. For students whose household education is below Grade 10, T2 increases senior English score by 0.215 SD, and final marks by 0.162 SD. The corresponding estimates are smaller among students from more-educated households. This pattern shows that school-based lessons and parental messages may be more valuable for students who receive less educational support at home.

# 8 Robustness

We show robustness of our main results on different estimation samples and sensitivity to controls.

**Full endline sample** Our primary estimation uses the matched sample of students surveyed at both baseline and endline. Table A21 shows the robustness with all the students present at the endline survey ($\sim$ 12.5k students). The results are largely consistent with the main estimates presented in Table 2. The positive effects on English test scores among Grades 9–10 students and on the confidence index remain statistically significant and similar in magnitude. The estimates for aspirations, gender attitudes, and socio-emotional learning are also comparable. However, the precision of the estimation is somewhat less. The T2 effect on final examination marks remains positive and statistically significant. Given the substantial inflow of new students into KGBVs during the intervention period, some attenuation in the

estimated effects is expected, but the overall pattern of results remains unchanged.

**Alternative controls** Figures 4 and 5 examine the robustness of the estimated effects on students' primary outcomes using four alternative sets of controls. Figure 4a presents the main specification from Table 2, which incorporates all baseline controls, whereas Figure 4b shows estimates without any controls. Figure 5a includes only student- and household-level controls, and Figure 5b includes only warden- and KGBV-level controls. The estimated treatment effects are consistent in magnitude across all four specifications, showing robustness to the selection of baseline controls.

**Attrition** Table A22 examines whether treatment assignment affected student attrition between baseline and endline. The pooled treatment reduced attrition by 3.2 percentage points relative to a control-group mean of 33%. This effect is driven primarily by T1, which reduced attrition by 5.5 percentage points, while the T2 estimate is small and statistically insignificant. The difference between the two treatment arms is marginally significant (p=0.057). These results indicate differential attrition among the T1. Baseline students who could not be surveyed at endline may have left the KGBV because of school dropout, transfer, or other reasons. Given the relatively high student turnover in the KGBV system, dropout and school transfer may account for a large share of the unmatched students. If so, the lower unmatched cases in the treatment group could reflect improved student retention. However, this interpretation remains suggestive and requires further investigation.

Table A23 reports inverse-probability-weighted (IPW) estimates that reweight observations by their estimated probability of remaining in the endline sample. The IPW estimates are remarkably similar to the main ITT results reported in Table 2. Across both cognitive and non-cognitive outcomes, the magnitudes and statistical significance of the treatment effects change little after reweighting. Specifically, the positive effects on English scores among Grades 9–10 students, confidence, and aspirations remain virtually unchanged. These results suggest that differential attrition is unlikely to bias the main findings.

# 9 Spillover Effects

This section examines whether the interventions targeted at students and parents generated spillover effects among individuals who were not directly exposed to the programme.

**Newly Admitted Grade 6 Students.** The student intervention, comprising the Life Skills and Foundational English curricula, was delivered to students in Grades 6–9. By the end of the intervention period, these students had progressed to Grades 7–10. Between October 2024 and October 2025, both treatment and control KGBVs admitted a new cohort of Grade 6 students. These newly admitted students did not directly receive the intervention. However,

students entering treatment KGBVs may have been indirectly exposed to the programme through interactions with older students who received the intervention or through changes in the broader school and hostel environment. We therefore examine whether treatment assignment generated spillover effects on cognitive outcomes, educational aspirations, and attitudes among the newly admitted Grade 6 students.

Appendix Table A31 reports effects on cognitive outcomes and educational aspirations. In the pooled specification, we find no statistically significant effects on final marks, English scores, Raven scores, or aspirations to study beyond Grade 12. We find suggestive evidence of a positive effect on students' perceived ability to obtain a university degree: students in treatment KGBVs are 5.4 percentage points more likely to report believing that they can obtain a university degree relative to students in control KGBVs ($p = 0.076$).

The treatment-arm-specific estimates reveal some differences across the two interventions. We find no statistically significant effects of T1 on the three cognitive outcomes or on aspirations to study beyond Grade 12. However, students in T1 KGBVs are 6.6 percentage points more likely to believe that they can obtain a university degree ($p = 0.081$), although this effect is statistically significant only at the 10 percent level. For T2, we find a positive effect on Raven scores of 0.258 standard deviations ($p = 0.036$) and suggestive evidence of a 7.6 percentage point increase in aspirations to study beyond Grade 12 ($p = 0.076$). The T1 and T2 effects on Raven scores are statistically different from one another ($p = 0.003$), while we cannot reject equality of the treatment effects for the remaining outcomes.

Appendix Table A32 examines spillover effects on socio-emotional learning outcomes. We find no statistically significant effects of the pooled treatment on the overall SEL index or on self-awareness, social awareness, self-management, and anxiety. The treatment-arm-specific estimates similarly provide little evidence of spillovers on socio-emotional outcomes. The only suggestive effect is observed for social awareness under T1, with an increase of 0.222 standard deviations ($p = 0.087$). However, the estimated effects of T1 and T2 are not statistically different from one another for any of the socio-emotional outcomes.

Table A33 reports spillover effects on the Gender Attitude Index and its components. The pooled treatment has a positive and statistically significant effect on attitudes toward education, increasing the standardized education-attitude measure by 0.325 standard deviations ($p = 0.001$). We find no statistically significant pooled effects on the overall Gender Attitude Index, employment attitudes, or equal-rights attitudes.

The treatment-arm-specific estimates show that the positive effect on education attitudes is present in both treatment arms. Education attitudes increase by 0.375 standard deviations under T1 ($p = 0.001$) and by 0.279 standard deviations under T2 ($p = 0.015$); the two treatment effects are not statistically different from one another ($p = 0.364$). T1 also produces

suggestive evidence of an improvement in equal-rights attitudes of 0.206 standard deviations ($p = 0.075$), although the pooled and T2 estimates are not statistically significant. While neither treatment arm individually has a statistically significant effect on the overall Gender Attitude Index, the estimated effects differ significantly across the two treatment arms ($p = 0.015$), with a negative point estimate for T1 and a positive point estimate for T2.

**Siblings of KGBV Hostel Students.** We additionally examine whether the interventions generated spillover effects on the siblings of students residing in KGBV hostels. Such spillovers may arise through communication and interactions between treated KGBV students and their siblings. In T2, an additional potential channel operates through parents, who directly received the parental IVR intervention and may subsequently transmit information, attitudes, or aspirations to other children in the household.

Table A34 reports spillover effects on siblings' cognitive outcomes, gender attitudes, general awareness, and educational aspirations. In the pooled specification, we find no statistically significant effects on Raven scores, ASER scores, gender attitudes, or university ability. The estimated effect on general awareness is positive but only marginally insignificant ($\beta = 0.226$, $p = 0.101$). In contrast, we find a positive and statistically significant effect on university aspirations: siblings in treatment households are 13.3 percentage points more likely to aspire to university education relative to siblings in the control group ($p = 0.001$).

The treatment-arm-specific estimates suggest that the increase in university aspirations is present under both interventions. T1 increases university aspirations by 15.5 percentage points ($p < 0.001$), while T2 increases them by 11.1 percentage points ($p = 0.024$). The two estimates are not statistically different from one another ($p = 0.298$), providing no evidence that the additional parental IVR component generated a differential effect on this outcome. We also find a positive effect of T1 on general awareness of 0.285 standard deviations ($p = 0.038$), while the corresponding T2 estimate is positive but not statistically significant. T1 additionally generates suggestive evidence of a 9.5 percentage point increase in siblings' perceived ability to attain university education ($p = 0.097$). We find no statistically significant effects of either treatment arm on Raven scores, ASER scores, or gender attitudes.

# 10 Policy implications and cost-effectiveness

The estimated effect of the lesson-plus-IVR intervention on the English performance of students in Grades 9–10 is 0.203 SD. Applying the conventional benchmark of 0.8 SD of learning per year proposed by Angrist et al. (2023b), this effect corresponds to approximately 0.25 learning-adjusted years of schooling (LAYS). In other words, the program generated learning equivalent to roughly one-quarter of a year of high-quality schooling.

The estimated program cost is approximately US$13.03 per T2 student. US$11.09 for the English and life-skills lessons and US$1.94 for the parental IVR component. The program therefore generated approximately 1.92 LAYS per US$100 spent.

The benchmarking exercise in Appendix Table 9 places the intervention in the middle of the distribution of comparable education programs. Its cost-effectiveness exceeds that of the para-teacher intervention studied by Eble et al. (2021), the after-school remedial program studied by Duflo et al. (2024), and the technology-based intervention studied by Muralidharan et al. (2019). It is less cost-effective, in terms of measured learning gains alone, than several low-cost remedial education, teacher-training, and mobile-learning interventions. This comparison nevertheless understates the program's broader benefits because LAYS captures only cognitive achievement. The intervention also generated improvements in educational aspirations, confidence, engagement, agency, and students' support networks, none of which enters the LAYS calculation. Moreover, the program is also intended to generate spillover effects among the siblings of the study students and the grade six students who were admitted to the KGBVs but did not attend the lessons. A more rigorous cost-effectiveness analysis might further improve the cost-effectiveness.

The policy relevance of the program also derives from its delivery model. The lessons were implemented through existing residential-school staff, while the relatively inexpensive IVR component provided a scalable mechanism for involving parents. However, cross-study comparisons should be interpreted cautiously because the interventions significantly differ in target populations and context. The table should therefore be viewed as a descriptive benchmarking exercise rather than a formal ranking of social returns.

# 11 Conclusion and Discussions

The results show a consistent pattern across students and parents. The intervention improved outcomes related to educational investment and future opportunities. Improvement in English skills, school exam performance, confidence, and educational aspirations support this claim. These improvements are accompanied by more educational engagement, lower perceived gender and family barriers, and better peer and warden relationships.

The student and parent results are reinforcing. Students become more confident and more likely to aspire to continue education beyond Grade 12. On the other hand, parents put greater importance on girls' education and perceived higher economic returns to their daughters' education. They are also more confident in their daughters' educational and professional success. T1 parents received no direct intervention. The positive parent effects in T1 suggest that these changes may have been transmitted indirectly through daughters exposed to the learning program. Thus, the program may have affected not only students' own beliefs

about their future but also household expectations regarding the returns to investing in girls' education.

The results also indicate that the combined intervention affected educational behavior as well as expectations. T2 students are more likely to attend school regularly and spend more time studying. They also report lower perceived family barriers to continuing education. These results are consistent with stronger academic effects under T2, especially for older students. They are also consistent with parental engagement reinforcing the student-level intervention.

The results also highlight the importance of the hostel environment. The program was delivered through existing wardens rather than external teachers. Students report better peer relationships and higher access to wardens for academic support. At the same time, parents report remarkably higher satisfaction with the hostel. They also report modest increases in communication with wardens and daughters. These findings indicate that the program may also have contributed to a more supportive environment for students. This is particularly important in residential schools.

The stronger effects among older students are noteworthy. Grades 9–10 students are closer to consequential decisions regarding secondary-school completion, higher education, and entry into the labor market. English proficiency may therefore have greater value for this group. Another possible explanation is that older students were better able to engage with the lesson content and translate the lessons into independent study.

A major challenge in implementing and evaluating programs among highly disadvantaged populations is maintaining a steady study cohort. The students enrolled in KGBVs come from some of the most socioeconomically disadvantaged and historically marginalized communities in India. Many face chronic poverty, seasonal migration, and early marriage. These factors contribute to frequent dropout or school transfers. This may cause higher student turnover in KGBVs than in traditional public schools. The high attrition observed in our sample highlights the importance of evaluating education interventions under the conditions in which they are ultimately intended to operate. Interventions that remain effective despite substantial student turnover are likely to be valuable for programs targeting the most disadvantaged populations.

# References


Afridi, F., Dinkelman, T., and Mahajan, K. (2018). Why are fewer married women joining the work force in rural india? a decomposition analysis over two decades. *Journal of Population Economics*, 31(3):783–818.

Alan, S. and Kubilay, E. (2025). Empowering adolescents to transform schools: Lessons from a behavioral targeting. *American Economic Review*, 115(2):365–407.

Albornoz, F., Anauati, M. V., Furman, M., Luzuriaga, M., Podesta, M. E., and Taylor, I. (2020). Training to teach science: experimental evidence from argentina. *The World Bank Economic Review*, 34(2):393–417.

Alix-García, J. and Knittel, C. R. (2025). Growing awareness: evaluating the impact of environmental education on attitudes, knowledge, and behavior. Technical report, National Bureau of Economic Research.

Andrew, A., Krutikova, S., Smarrelli, G., and Verma, H. (2022). Gender norms, violence and adolescent girls' trajectories: evidence from a field experiment in india. Technical report, IFS Working Papers.

Angrist, N., Ainomugisha, M., Bathena, S. P., Bergman, P., Crossley, C., Cullen, C., Letsomo, T., Matsheng, M., Panti, R. M., Sabarwal, S., et al. (2023a). Building resilient education systems: Evidence from large-scale randomized trials in five countries. Technical report, National Bureau of Economic Research.

Angrist, N., Evans, D. K., Filmer, D., Glennerster, R., Rogers, H., and Sabarwal, S. (2023b). How to improve education outcomes most efficiently?

Anthony, C. J., Elliott, S. N., DiPerna, J. C., and Lei, P.-W. (2020). The ssis sel brief scales–student form: initial development and validation. *School Psychology*, 35(4):277.

Appadurai, A. (2007). The capacity to aspire: Culture and the terms of recognition. *Cultural politics in a global age: Uncertainty, solidarity and innovation*, page 29.

Asadullah, M. N., Chaudhury, N., and Josh, S. R. A.-Z. (2009). Secondary school madrasas in bangladesh: incidence, quality, and implications for reform. *Human Development Sector, The World Bank. Retrieved from: http://siteresources. worldbank. org/EXTDEVDIALOGUE/Resources/BangladeshMadrasa ReportFinal. pdf.*

Ayonrinde, O. T., Ayonrinde, O. A., Adams, L. A., Sanfilippo, F. M., O'Sullivan, T. A., Robinson, M., Oddy, W. H., and Olynyk, J. K. (2020). The relationship between abdominal pain and emotional wellbeing in children and adolescents in the raine study. *Scientific reports*, 10(1):1646.

Azam, M., Chin, A., and Prakash, N. (2013). The returns to english-language skills in india. *Economic Development and Cultural Change*, 61(2):335–367.

Bandiera, O., Buehren, N., Burgess, R., Goldstein, M., Gulesci, S., Rasul, I., and Sulaiman, M. (2020). Women's empowerment in action: evidence from a randomized control trial in africa. *American Economic Journal: Applied Economics*, 12(1):210–259.

Banerjee, A. V., Cole, S., Duflo, E., and Linden, L. (2007). Remedying education: Evidence from two randomized experiments in india. *The quarterly journal of economics*, 122(3):1235–1264.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *science*, 335(6068):582–586.

Bharti, N. K., Blakeslee, D., Walter, T. F., Malik, S., and Naeem, T. (2025). Clean air in the classroom: Environmental inputs and human capital formation.

Bharti, N. K., Malik, S., Mukhopadhyay, A., and Prakash, N. (2026). *Science on the Move: How Experiential Pedagogy Shapes Human Capital*. Munich Society for the Promotion of Economic Research-CESifo GmbH.

Bhattacharya, S., Constantino, S., Mishra, N., Prakash, N., Sabarwal, S., Samaddar, D., and Sherif, R. (2026). Intergenerational spillovers of environmental attitudes and behaviors.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160.

Census India (2011). Census tables. https://censusindia.gov.in/census.website/data/census-tables. Ministry of Home Affairs, Government of India. Accessed 15 February 2025.

Choudhary, A. and Singh, A. (2017). Are daughters like mothers: evidence on intergenerational educational mobility among young females in india. *Social Indicators Research*, 133(2):601–621.

Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and aspirations failure. *The Economic Journal*, 126(590):165–188.

Deininger, K., Goyal, A., and Nagarajan, H. (2013). Women's inheritance rights and intergenerational transmission of resources in india. *Journal of Human Resources*, 48(1):114–141.

Desai, S. and Kulkarni, V. (2008). Changing educational inequalities in india in the context of affirmative action. *Demography*, 45(2):245–270.

Dhar, D., Jain, T., and Jayachandran, S. (2019). Intergenerational transmission of gender attitudes: Evidence from india. *The Journal of Development Studies*, 55(12):2572–2592.

Dhar, D., Jain, T., and Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American economic review*, 112(3):899–927.

DoSJE (2024). Handbook on social welfare statistics. *Department of Social Justice & Empowerment Statistics Division, Ministry of Social Justice & Empowerment, Government of India*. Accessed 05 December 2025.

Duflo, A., Kiessel, J., and Lucas, A. M. (2024). Experimental evidence on four policies to increase learning at scale. *The Economic Journal*, 134(661):1985–2008.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic*

*literature*, 50(4):1051–1079.

Duflo, E., Dupas, P., and Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from kenyan primary schools. *Journal of public Economics*, 123:92–110.

Dunn, D. (1993). Gender inequality in education and employment in the scheduled castes and tribes of india. *Population Research and Policy Review*, 12(1):53–70.

Eble, A., Frost, C., Camara, A., Bouy, B., Bah, M., Sivaraman, M., Hsieh, P.-T. J., Jayanty, C., Brady, T., Gawron, P., et al. (2021). How much can we remedy very low learning levels in rural parts of low-income countries? impact and generalizability of a multi-pronged para-teacher intervention from a cluster-randomized trial in the gambia. *Journal of Development Economics*, 148:102539.

Edmonds, E., Feigenberg, B., and Leight, J. (2023). Advancing the agency of adolescent girls. *Review of Economics and Statistics*, 105(4):852–866.

Ferrara, E. L., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Field, E. and Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of political Economy*, 116(5):881–930.

Field, E., Jayachandran, S., and Pande, R. (2010). Do traditional institutions constrain female entrepreneurship? a field experiment on business training in india. *American Economic Review*, 100(2):125–129.

Field, E., Pande, R., Rigol, N., Schaner, S., and Troyer Moore, C. (2021). On her own account: How strengthening women's financial control impacts labor supply and gender norms. *American Economic Review*, 111(7):2342–2375.

Ganicho, A., Magalhães, S., Limpo, T., Lemos, M. S., and Alves, D. (2025). The ssis sel brief scales–student form: Examining the structural validity of a portuguese version. In *Child & Youth Care Forum*, volume 54, pages 245–264. Springer.

Ghatak, D., Sahoo, S., Sarkar, S., and Sharma, V. (2024). Who eats last? intra-household gender inequality in food allocation among children in educationally backward areas of india. *Population Studies*, 78(1):63–77.

Giacobino, H., Huillery, E., Michel, B., and Sage, M. (2024). Schoolgirls, not brides: Education as a shield against child marriage. *American Economic Journal: Applied Economics*, 16(4):109–143.

Goel, S. and Husain, Z. (2018). Gender, caste, and education in india: A cohort-wise study of drop-out from schools. *Research in Social Stratification and Mobility*, 58:54–68.

Gunawardena, N., Kurotani, K., Indrawansa, S., Nonaka, D., Mizoue, T., and Samarasinghe, D. (2016). School-based intervention to enable school children to act as change agents on weight, physical activity and diet of their mothers: a cluster randomized controlled trial. *International Journal of Behavioral Nutrition and Physical Activity*, 13(1):45.

Heckman, J. J. and Mosso, S. (2014). The economics of human development and social

mobility. *Annu. Rev. Econ.*, 6(1):689–733.

ILO (2024). Ilostat: Indicators and data tools. https://ilostat.ilo.org/data/. Accessed 15 December 2025.

Jain, T., Maitra, P., and Mani, S. (2019). Barriers to skill acquisition: Evidence from english training in india. *World Development*, 114:314–325.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual review of economics*, 7(1):63–88.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review*, 69(3):576–595.

Jensen, R. and Oster, E. (2009). The power of tv: Cable television and women's status in india. *The Quarterly Journal of Economics*, 124(3):1057–1094.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., and Borghans, L. (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success.

Kishor, S. and Gupta, K. (2004). Women's empowerment in india and its states: evidence from the nfhs. *Economic and Political weekly*, pages 694–712.

Klasen, S. and Pieters, J. (2015). What explains the stagnation of female labor force participation in urban india? *The World Bank Economic Review*, 29(3):449–478.

La Ferrara, E. (2019). Presidential address: Aspirations, social norms, and development. *Journal of the European Economic Association*, 17(6):1687–1722.

Lam, S.-f., Wong, B. P. H., Yang, H., and Liu, Y. (2012). Understanding student engagement with a contextual model. In Christenson, S. L., Reschly, A. L., and Wylie, C., editors, *Handbook of Research on Student Engagement*, chapter 19, pages 403–419. Springer, Boston, MA.

Loyalka, P., Mistree, D., Fairlie, R., and Khanna, S. (2025). Job training, english language skills, and employability: Evidence from an experiment in urban india. *Economic Development and Cultural Change*, 73(4):000–000.

Luo, R., Shi, Y., Zhang, L., Liu, C., Rozelle, S., and Sharbono, B. (2009). Malnutrition in china's rural boarding schools: The case of primary schools in shaanxi province. *Asia Pacific Journal of Education*, 29(4):481–501.

Maertens, A. (2013). Social norms and aspirations: age of marriage and education in rural india. *World Development*, 47:1–15.

Mitra, A. (2008). The status of women among the scheduled tribes in india. *The Journal of Socio-Economics*, 37(3):1202–1217.

Muralidharan, K., Singh, A., and Ganimian, A. J. (2019). Disrupting education? experimental evidence on technology-aided instruction in india. *American Economic Review*, 109(4):1426–1460.

Nourani, V., Ashraf, N., and Banerjee, A. (2023). Learning to teach by learning to learn.

Porecha, M. (2024). Lack of transparency in spending, weak infrastructure, safety concerns: Ncert report highlights issues in kgbv schools. *The Hindu*. Accessed 05 December 2025.

Raghavendra, R. (2020). Literacy and health status of scheduled castes in india. *Contemporary Voice of Dalit*, 12(1):97–110.

Rao, N. (2014). Caste, kinship, and life course: Rethinking women's work and agency in rural south india. *Feminist Economics*, 20(3):78–102.

Ray, D. (2006). Aspirations, poverty, and economic change. *Understanding poverty*, 1(2006):409–421.

Sabates, R., Rose, P., Alcott, B., and Delprato, M. (2021). Assessing cost-effectiveness with equity of a programme targeting marginalised girls in secondary schools in tanzania. *Journal of Development Effectiveness*, 13(1):28–46.

Saravanakumar, S. and Palanisamy, M. (2013). Status of primary education of scheduled caste children. *International Journal of Research in Humanities and Social Sciences*, 1(8):54–61.

Sheehan, P., Sweeny, K., Rasmussen, B., Wils, A., Friedman, H. S., Mahon, J., Patton, G. C., Sawyer, S. M., Howard, E., Symons, J., et al. (2017). Building the foundations for sustainable development: a case for global investment in the capabilities of adolescents. *The Lancet*, 390(10104):1792–1806.

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., and Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? two meta-analyses. *Psychological science*, 29(4):549–571.

Song, Y. (2020). Four decades of compulsory education in china: Achievements, experience and challenges. *Center for International Knowledge on Development. http://archive. cikd. org/fileCache/pdf/F/Four Decades of Compulsory Education in China Achievements Experience and Challenges. pdf.*

Starr, K. J. and Sahgal, N. (2021). Measuring caste in india. *Pew Research Center (Decoded), June*, 29.

Steinberg, L. D. (2014). *Age of opportunity: Lessons from the new science of adolescence*. Houghton Mifflin Harcourt.

Sunder, N. (2015). Marriage age, social status and intergenerational effects in uganda. social status and intergenerational effects in uganda. *November*, 29:2015.

Thakur, A. (2020). Literacy rate for muslims worse than sc/sts. *The Times of India*. Accessed 15 February 2025.

Wang, A., Medina, A., Luo, R., Shi, Y., and Yue, A. (2016). To board or not to board: Evidence from nutrition, health and education outcomes of students in rural china. *China & World Economy*, 24(3):52–66.

White, K. S. and Farrell, A. D. (2006). Anxiety and psychosocial stress as predictors of headache and abdominal pain in urban early adolescents. *Journal of pediatric psychology*, 31(6):582–596.

# 12 Main Figures

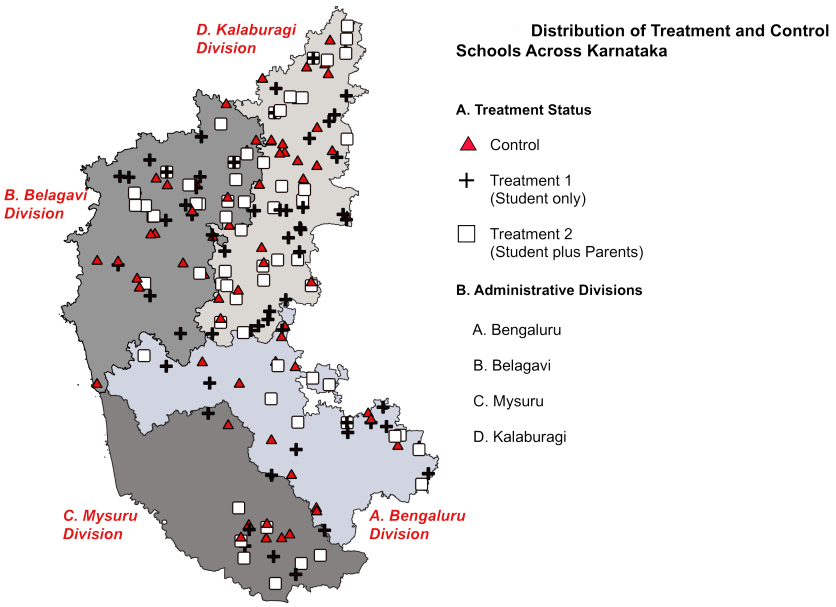



Figure 2: Geographic distribution of participating KGBVs by treatment assignment

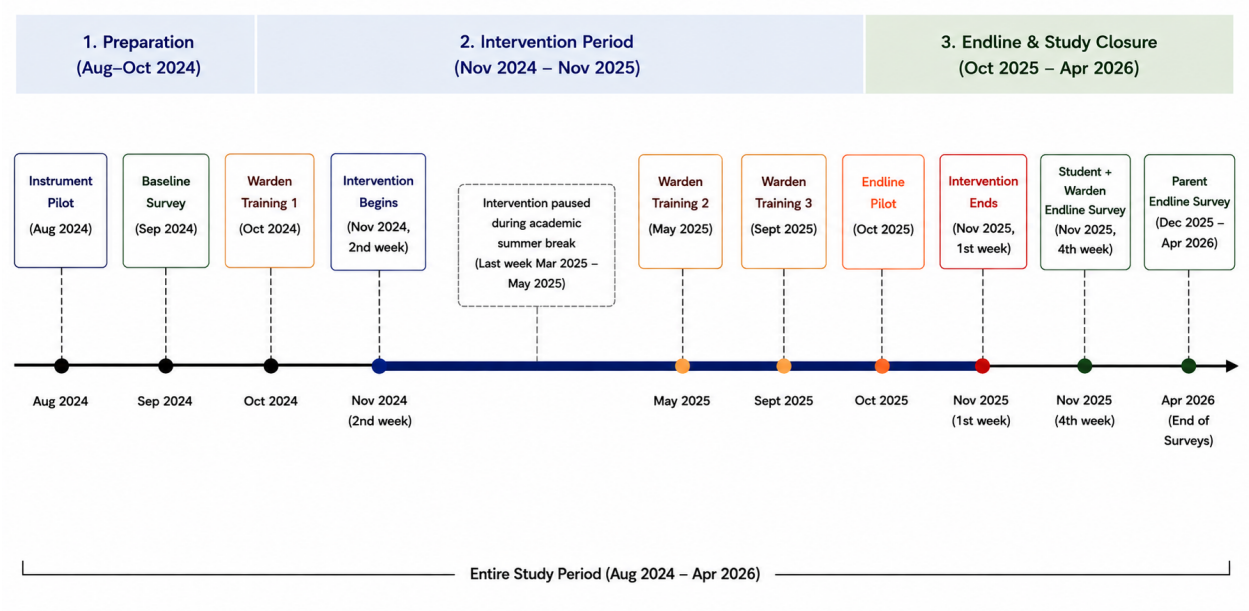
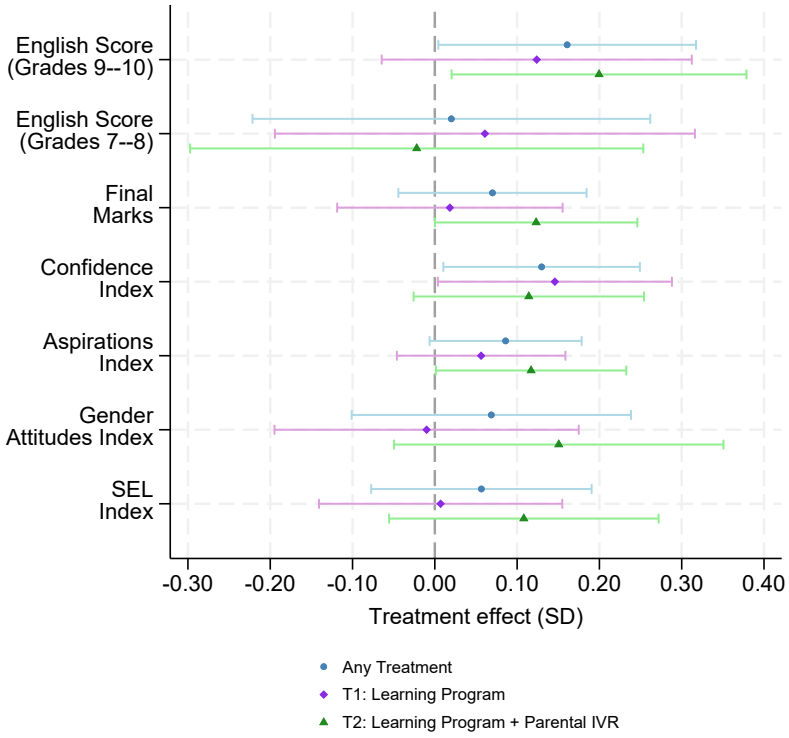
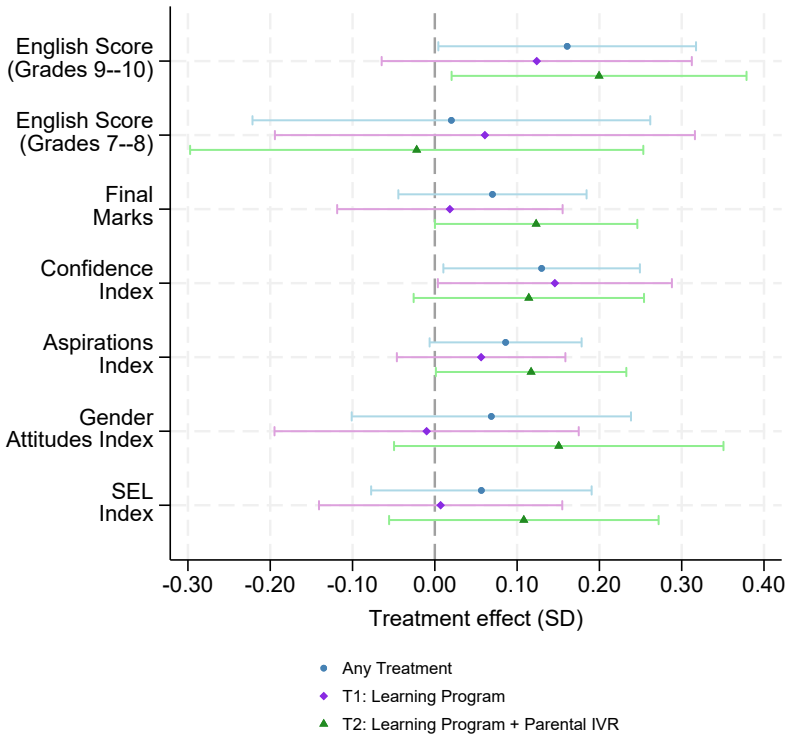



Figure 3: Study Timeline

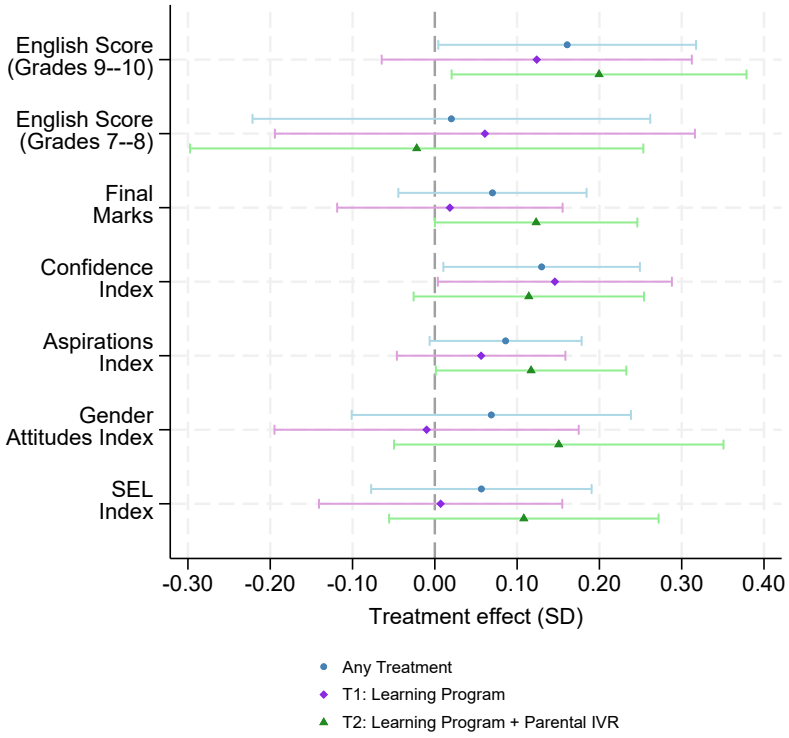



(a) Primary student outcomes with all basic controls

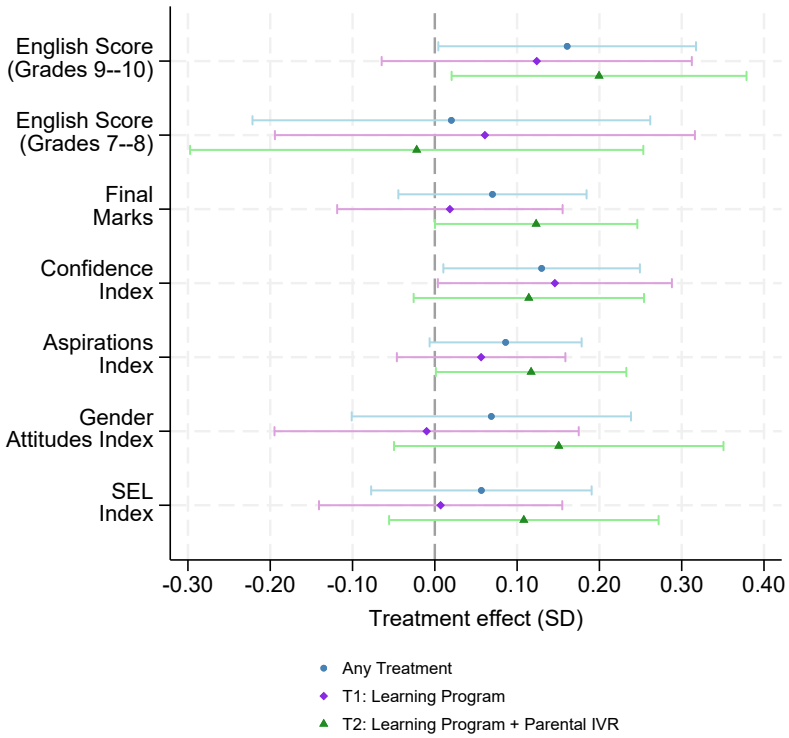



(b) Primary student outcomes with no controls

Figure 4: Treatment Effects on Primary Outcomes




(a) Primary student outcomes with basic student controls




(b) Primary student outcomes with Warden & KGBV Hostel controls

Figure 5: Treatment Effects on Primary Outcomes

# 13 Main Tables

Table 1: Effects on Warden Delivery of Structured Lessons

| | (1) Self-Awareness | (2) Social Awareness | (3) Life Skills | (4) Health and Hygiene | (5) Gender Sensitivity |
|---|---|---|---|---|---|
| Any treatment | 0.293*** | 0.333*** | 0.335*** | 0.265*** | 0.348*** |
| | (0.056) | (0.055) | (0.058) | (0.055) | (0.061) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| T1 (Learning Program) | 0.273*** | 0.309*** | 0.306*** | 0.243*** | 0.301*** |
| | (0.060) | (0.059) | (0.063) | (0.057) | (0.067) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| T2 (Learning Program + Parental IVR) | 0.315*** | 0.359*** | 0.366*** | 0.289*** | 0.398*** |
| | (0.058) | (0.057) | (0.060) | (0.058) | (0.063) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| p-value: T1 = T2 | 0.257 | 0.187 | 0.140 | 0.194 | 0.026 |
| Observations | 195 | 195 | 195 | 195 | 195 |
| Control Mean | 0.697 | 0.652 | 0.652 | 0.727 | 0.621 |
| Basic controls | Yes | Yes | Yes | Yes | Yes |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on wardens' delivery of structured lessons across five content domains. Each outcome is an indicator equal to one if the warden reported conducting lessons in the relevant domain. The pooled specification combines the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Regressions include basic warden and KGBV-level controls and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, and $^{***}$ $p < 0.01$.

Table 2: Effects on **Primary Outcomes**

| | **Cognitive Outcomes** | | | **Non-cognitive Outcomes** | | | |
|---|---|---|---|---|---|---|---|
| | (1) English Score (Grades 9–10) | (2) English Score (Grades 7–8) | (3) Final Marks | (4) Confidence Index | (5) Aspirations Index | (6) Gender Attitudes Index | (7) SEL Index |
| Any treatment | 0.161** | 0.020 | 0.070 | 0.130** | 0.086* | 0.069 | 0.057 |
| | (0.079) | (0.122) | (0.058) | (0.061) | (0.047) | (0.086) | (0.068) |
| | [0.044] | [0.869] | [0.229] | [0.033] | [0.068] | [0.426] | [0.406] |
| T1 (Learning Program) | 0.124 | 0.061 | 0.018 | 0.146** | 0.056 | -0.010 | 0.007 |
| | (0.096) | (0.129) | (0.069) | (0.072) | (0.052) | (0.094) | (0.075) |
| | [0.196] | [0.638] | [0.792] | [0.044] | [0.280] | [0.915] | [0.925] |
| T2 (Learning Program + Parental IVR) | 0.200** | -0.022 | 0.123** | 0.114 | 0.117** | 0.151 | 0.108 |
| | (0.091) | (0.139) | (0.062) | (0.071) | (0.059) | (0.102) | (0.083) |
| | [0.029] | [0.875] | [0.050] | [0.109] | [0.048] | [0.139] | [0.194] |
| p-value: T1 = T2 | 0.443 | 0.466 | 0.110 | 0.679 | 0.313 | 0.100 | 0.222 |
| Observations | 4621 | 4621 | 8730 | 6867 | 9243 | 9243 | 9243 |
| Control Mean | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 198 | 134 | 202 | 199 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on the primary student outcomes. The first specification pools the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Outcomes are standardized relative to the control group. All regressions include the same basic controls, as well as strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses. Conventional p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 3: Effects on Students' Reported Agency

| | (1) Menstruation Openness | (2) Shares Emotions Often | (3) Preferred Marriage Age | (4) Marriage Agency Index | (5) Locus of Control Index |
|---|---|---|---|---|---|
| Any treatment | 0.043*** | 0.044** | 0.118 | 0.123** | 0.090* |
| | (0.016) | (0.022) | (0.125) | (0.061) | (0.052) |
| | [0.006] | [0.047] | [0.347] | [0.046] | [0.085] |
| T1 (Learning Program) | 0.053*** | 0.032 | -0.023 | 0.090 | 0.088 |
| | (0.018) | (0.025) | (0.128) | (0.068) | (0.059) |
| | [0.004] | [0.209] | [0.855] | [0.185] | [0.137] |
| T2 (Learning Program + Parental IVR) | 0.033* | 0.057** | 0.277* | 0.157** | 0.092 |
| | (0.018) | (0.026) | (0.163) | (0.075) | (0.059) |
| | [0.059] | [0.032] | [0.090] | [0.037] | [0.123] |
| $p$-value: T1 = T2 | 0.263 | 0.338 | 0.044 | 0.367 | 0.947 |
| Observations | 6,720 | 8,893 | 3,343 | 9,243 | 9,243 |
| Control Mean | 0.841 | 0.598 | 22.640 | -0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 199 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' reported agency and related outcomes. Columns (1) and (2) report indicators for openness about menstruation and frequently sharing emotions. Column (3) reports students' preferred age at marriage in years. Columns (4) and (5) report standardized indices of marriage agency and locus of control. The pooled specification combines the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Regressions include the same basic controls and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 4: Effects on Perceived Barriers to Education and Self-Efficacy

| | (1) Gender Barriers | (2) Family Barriers | (3) Caste Barriers | (4) Economic Barriers | (5) Self-Efficacy |
|---|---|---|---|---|---|
| Any treatment | -0.058* | -0.033 | -0.039 | 0.003 | 0.014 |
| | (0.030) | (0.031) | (0.028) | (0.033) | (0.018) |
| | [0.053] | [0.288] | [0.166] | [0.930] | [0.448] |
| T1 (Learning Program) | -0.036 | 0.004 | -0.028 | 0.024 | -0.001 |
| | (0.034) | (0.035) | (0.032) | (0.036) | (0.021) |
| | [0.279] | [0.915] | [0.379] | [0.507] | [0.973] |
| T2 (Learning Program + Parental IVR) | -0.080** | -0.071* | -0.049 | -0.019 | 0.029 |
| | (0.038) | (0.036) | (0.035) | (0.041) | (0.021) |
| | [0.035] | [0.050] | [0.159] | [0.649] | [0.180] |
| $p$-value: T1 = T2 | 0.274 | 0.047 | 0.582 | 0.292 | 0.185 |
| Observations | 8,960 | 8,990 | 8,970 | 8,996 | 9,038 |
| Control Mean | 0.500 | 0.447 | 0.393 | 0.523 | 0.751 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' perceived barriers to education and self-efficacy. Columns (1)–(4) report indicators for whether students perceive gender, family, caste, and economic barriers to continuing their education. Column (5) reports the self-efficacy measure. The pooled specification combines the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. All regressions include the same basic controls along with a social desirability bias index and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, and $^{***}$ $p < 0.01$.

Table 5: Effects on Education Engagement

| | (1) Never Missed School | (2) Weekday Study Hours (Hostel) | (3) Weekly Study Hours (Hostel) | (4) Weekly Study Hours (Home) |
|---|---|---|---|---|
| Any treatment | 0.050* | 0.296 | 0.468 | 0.178 |
| | (0.027) | (0.379) | (0.504) | (0.462) |
| | [0.067] | [0.436] | [0.354] | [0.701] |
| T1 (Learning Program) | 0.024 | -0.357 | -0.275 | 0.023 |
| | (0.029) | (0.477) | (0.646) | (0.538) |
| | [0.412] | [0.454] | [0.670] | [0.967] |
| T2 (Learning Program + Parental IVR) | 0.078** | 0.974** | 1.239** | 0.339 |
| | (0.032) | (0.440) | (0.554) | (0.527) |
| | [0.016] | [0.028] | [0.027] | [0.521] |
| $p$-value: T1 = T2 | 0.068 | 0.013 | 0.025 | 0.553 |
| Observations | 8,856 | 9,110 | 9,161 | 8,935 |
| Control Mean | 0.616 | 22.450 | 28.742 | 16.656 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' education engagement. The sample includes the students matched between the baseline and endline surveys. The pooled specification combines the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Column (1) is an indicator for never missing school during last week. Columns (2)–(4) report study hours in their original units. All regressions include basic controls and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 6: Effects on Students' Support Network

| | (1) Peer Relationship Index | (2) Student–Warden Relationship Index | (3) Can Approach Warden for Study Help | (4) Calls Family at Least Weekly | (5) Comfortable Sharing Hostel Experiences |
|---|---|---|---|---|---|
| Any treatment | 0.135*** | 0.100 | 0.056** | 0.043 | 0.032 |
| | (0.049) | (0.066) | (0.024) | (0.038) | (0.025) |
| | [0.006] | [0.132] | [0.019] | [0.258] | [0.203] |
| T1 (Learning Program) | 0.127** | 0.065 | 0.033 | 0.015 | 0.014 |
| | (0.062) | (0.080) | (0.028) | (0.041) | (0.027) |
| | [0.042] | [0.413] | [0.234] | [0.723] | [0.596] |
| T2 (Learning Program + Parental IVR) | 0.144** | 0.137* | 0.079*** | 0.072 | 0.051* |
| | (0.061) | (0.075) | (0.027) | (0.048) | (0.030) |
| | [0.019] | [0.071] | [0.004] | [0.137] | [0.091] |
| $p$-value: T1 = T2 | 0.821 | 0.385 | 0.107 | 0.242 | 0.175 |
| Observations | 9,243 | 9,243 | 8,976 | 8,956 | 8,992 |
| Control Mean | -0.000 | 0.000 | 0.706 | 0.560 | 0.704 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' support networks and relationships. Columns (1) and (2) report standardized indices of peer relationships and student–warden relationships. Columns (3)–(5) report indicators for whether a student can approach the warden for study-related help, calls her family at least weekly, and feels comfortable sharing her hostel experiences. The pooled specification combines the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. All regressions include the same basic controls along with a social disability bias index and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 7: Effects on Parents' Education and Gender Attitudes

| | (1) Importance of Education | (2) Employment Attitudes | (3) Gender Attitude | (4) Social Norms Attitude | (5) Equal Rights Attitude |
|---|---|---|---|---|---|
| Any treatment | 0.355*** | 0.196** | 0.140 | 0.104 | -0.048 |
| | (0.071) | (0.083) | (0.093) | (0.099) | (0.086) |
| | [0.000] | [0.019] | [0.136] | [0.297] | [0.575] |
| T1 (Learning Program) | 0.434*** | 0.224** | 0.178 | 0.154 | -0.002 |
| | (0.082) | (0.098) | (0.113) | (0.106) | (0.099) |
| | [0.000] | [0.023] | [0.116] | [0.148] | [0.983] |
| T2 (Learning Program + Parent IVR) | 0.277*** | 0.168* | 0.101 | 0.054 | -0.094 |
| | (0.088) | (0.094) | (0.109) | (0.130) | (0.105) |
| | [0.002] | [0.077] | [0.354] | [0.678] | [0.373] |
| p-value: T1 = T2 | 0.102 | 0.566 | 0.524 | 0.443 | 0.410 |
| Observations | 4216 | 4216 | 4216 | 4216 | 4216 |
| Control Mean | -0.000 | 0.000 | -0.000 | 0.000 | -0.000 |
| Parent Controls | Yes | Yes | Yes | Yes | Yes |
| KGBV Controls | Yes | Yes | Yes | Yes | Yes |
| Strata FE | Yes | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 203 | 203 | 203 |

*Notes:* The table reports intention-to-treat estimates of the effects of the intervention on parents' attitudes toward girls' education, employment, gender norms, social norms, and equal rights. The first specification pools Treatment 1 and Treatment 2, while the subsequent specifications report treatment-arm-specific effects. Treatment 1 consists of the student-level learning programme, whereas Treatment 2 combines the student-level learning programme with the parent IVR intervention. All outcomes are standardized indices. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Robust standard errors clustered at the KGBV level are reported in parentheses, and p-values are reported in square brackets. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, and $^{***}$ $p < 0.01$.

Table 8: Effects on Parents' Confidence in their Daughters

| | (1) Confidence in Educational Attainment | (2) Confidence in Professional Achievement | (3) Expected Earnings Grade 8 | (4) Expected Earnings Grade 10 | (5) Expected Earnings Grade 12 | (6) Expected Earnings College |
|---|---|---|---|---|---|---|
| Any treatment | 0.123*** | 0.111*** | 0.098 | 0.088** | 0.067** | 0.109** |
| | (0.028) | (0.026) | (0.063) | (0.044) | (0.034) | (0.044) |
| | [0.000] | [0.000] | [0.119] | [0.046] | [0.047] | [0.014] |
| T1 (Learning Program) | 0.158*** | 0.113*** | 0.146* | 0.110** | 0.098** | 0.148*** |
| | (0.035) | (0.033) | (0.074) | (0.054) | (0.044) | (0.053) |
| | [0.000] | [0.001] | [0.051] | [0.041] | [0.026] | [0.006] |
| T2 (Learning Program + Parent IVR) | 0.089*** | 0.109*** | 0.050 | 0.065 | 0.037 | 0.070 |
| | (0.031) | (0.029) | (0.074) | (0.048) | (0.036) | (0.048) |
| | [0.004] | [0.000] | [0.499] | [0.180] | [0.312] | [0.147] |
| p-value: T1 = T2 | 0.062 | 0.913 | 0.223 | 0.390 | 0.161 | 0.125 |
| Observations | 4216 | 4216 | 2978 | 3309 | 3636 | 4044 |
| Control Mean | 0.143 | 0.162 | 8.496 | 8.927 | 9.306 | 9.701 |
| Parent Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| KGBV Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Strata FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 183 | 191 | 196 | 203 |

*Notes:* The table reports intention-to-treat estimates of the effects of the intervention on parents' confidence in their daughters' educational attainment and professional achievement, as well as parents' expectations of daughters' earnings at different levels of education. The first specification pools Treatment 1 and Treatment 2, while the subsequent specifications report treatment-arm-specific effects. Treatment 1 consists of the student-level learning programme, whereas Treatment 2 combines the student-level learning programme with the parent IVR intervention. The confidence outcomes are binary indicators equal to one when parents report high confidence that their daughters will achieve the specified educational or professional outcome. Expected earnings are reported in logarithmic units. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Robust standard errors clustered at the KGBV level are reported in parentheses, and p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 9: Benchmarking Learning Gains and Cost-Effectiveness

| Study | Intervention | Country | Effect Size | LAYS | Cost per Student | LAYS per US$100 |
|---|---|---|---|---|---|---|
| Duflo et al. (2024) | After-school remedial education | Ghana | 0.15 | 0.19 | 132.00 | 0.14 |
| Eble et al. (2021) | Para-teacher intervention | Gambia | 3.23 | 4.04 | 724.00 | 0.56 |
| Muralidharan et al. (2019) | Technology-aided instruction | India | 0.37 | 0.46 | 67.50 | 0.69 |
| **Our Study** | **English lessons + IVR** | **India** | **0.19** | **0.24** | **13.03** | **1.84** |
| Banerjee et al. (2007) | Computer-aided learning | India | 0.35 | 0.44 | 15.18 | 2.88 |
| Angrist et al. (2023a) | Phone-call and SMS tutorials | Multi-country | 0.33 | 0.41 | 11.00 | 3.72 |
| Bharti et al. (2025) | Classroom air purifiers | Pakistan | 0.15 | 0.19 | 4.72 | 3.97 |
| Duflo et al. (2015) | Contract teachers | Kenya | 0.20 | 0.25 | 5.47 | 4.48 |
| Sabates et al. (2021) | Girls' education programme | Tanzania | 0.60 | 0.75 | 15.00 | 5.02 |
| Albornoz et al. (2020) | Teacher training | Argentina | 0.65 | 0.81 | 14.70 | 5.50 |
| Bharti et al. (2026) | Mobile science laboratory | India | 0.32 | 0.40 | 5.56 | 7.19 |
| Banerjee et al. (2007) | Remedial education | India | 0.14 | 0.18 | 2.35 | 7.45 |
| Nourani et al. (2023) | Teacher training | Uganda | 0.51 | 0.64 | 6.62 | 9.63 |

*Notes:* Effect sizes are reported in standard-deviation units. Learning-adjusted years of schooling (LAYS) are calculated by converting each estimated learning effect using a benchmark of 0.8 SD of learning per year of schooling (Angrist et al., 2023b). Cost-effectiveness is measured as LAYS generated per US$100 spent per student. Costs are expressed in US dollars. Studies are ordered by estimated LAYS generated per US$100.

# Appendix

Table A1: Sample distribution across treatment and control groups

| | Treatment 1 | Treatment 2 | Control | Total |
|---|---|---|---|---|
| **Panel A: KGBV** | | | | |
| Randomized | 69 | 66 | 70 | 205 |
| Baseline | 68 | 65 | 70 | 203 |
| Endline | 68 | 65 | 69 | 202 |
| **Panel B: Student** | | | | |
| Baseline | 4,339 | 4,444 | 4,600 | 13,383 |
| Endline | 5,159 | 5,090 | 5,111 | 15,360 |
| Matched | 3,141 | 3,022 | 3,080 | 9,243 |
| **Panel C: Parent** | | | | |
| Baseline | 1,570 | 1,592 | 1,640 | 4,802 |
| Endline | 1,432 | 1,413 | 1,371 | 4,216 |
| **Panel D: Warden** | | | | |
| Baseline | 68 | 65 | 69 | 202 |
| Endline | 68 | 65 | 69 | 202 |
| Matched | 54 | 54 | 50 | 158 |

Table A2: Balance of Baseline Characteristics (Full Baseline Student Sample)

| Variable | (1) Control Mean | (2) T1: Mean | (3) T2: Mean | (4) T1 – Control | (5) T2 – Control |
|---|---|---|---|---|---|
| **Panel A: Student Characteristics and Baseline Outcomes** | | | | | |
| Age (years) | 13.382 | 13.381 | 13.538 | 0.021 | 0.148** |
| | (0.096) | (0.097) | (0.098) | | |
| Share of students in Grade 6 | 0.181 | 0.181 | 0.158 | -0.005 | -0.023 |
| | (0.020) | (0.018) | (0.018) | | |
| Share of students in Grade 7 | 0.190 | 0.194 | 0.183 | -0.001 | -0.006 |
| | (0.019) | (0.018) | (0.020) | | |
| Share of students in Grade 8 | 0.218 | 0.214 | 0.220 | -0.011 | 0.002 |
| | (0.020) | (0.019) | (0.021) | | |
| Share of students in Grade 9 | 0.411 | 0.411 | 0.439 | 0.017 | 0.027 |
| | (0.051) | (0.051) | (0.051) | | |
| Years enrolled in KGBV | 1.879 | 1.781 | 1.888 | -0.095 | 0.010 |
| | (0.078) | (0.073) | (0.060) | | |
| Days absent from school last week | 1.065 | 0.922 | 1.214 | -0.153 | 0.156 |
| | (0.080) | (0.097) | (0.093) | | |
| Daily study hours in hostel | 3.212 | 3.129 | 3.109 | -0.072 | -0.107 |
| | (0.055) | (0.076) | (0.064) | | |
| Total cognitive test score | 11.234 | 12.070 | 11.418 | 0.830 | 0.205 |
| | (0.284) | (0.449) | (0.370) | | |
| English test score | 4.809 | 5.226 | 4.894 | 0.416* | 0.094 |
| | (0.138) | (0.212) | (0.173) | | |
| Previous examination marks (%) | 66.012 | 66.594 | 68.251 | 0.662 | 2.023 |
| | (1.112) | (1.247) | (1.182) | | |
| ASER score | 2.451 | 2.580 | 2.641 | 0.140 | 0.213 |
| | (0.101) | (0.091) | (0.120) | | |
| Perceived barriers index | -0.041 | -0.117 | -0.021 | -0.087 | 0.026 |
| | (0.061) | (0.059) | (0.070) | | |
| Gender attitudes index | 0.023 | 0.106 | 0.032 | 0.093 | 0.011 |
| | (0.054) | (0.072) | (0.053) | | |
| Aspirations index | 0.031 | 0.008 | -0.036 | -0.016 | -0.068 |
| | (0.050) | (0.051) | (0.053) | | |
| Joint F-test $p$-value | | | | 0.320 | 0.144 |
| Number of clusters | 70 | 68 | 65 | | |
| **Panel B: Household Characteristics** | | | | | |
| Number of household members | 6.597 | 6.337 | 6.556 | -0.275** | -0.070 |
| | (0.114) | (0.094) | (0.118) | | |
| Number of siblings | 2.671 | 2.627 | 2.710 | -0.055 | 0.036 |
| | (0.066) | (0.074) | (0.077) | | |
| First-born child | 0.463 | 0.448 | 0.433 | -0.015 | -0.029 |
| | (0.013) | (0.014) | (0.014) | | |
| Father's education (years) | 6.791 | 6.651 | 6.964 | -0.136 | 0.213 |
| | (0.211) | (0.212) | (0.208) | | |
| Mother engaged in farming | 0.613 | 0.676 | 0.599 | 0.064** | -0.010 |
| | (0.020) | (0.023) | (0.024) | | |
| Father engaged in farming | 0.509 | 0.577 | 0.502 | 0.068** | -0.007 |
| | (0.024) | (0.026) | (0.023) | | |
| Number of rooms in household | 2.590 | 2.527 | 2.584 | -0.066 | -0.008 |
| | (0.060) | (0.052) | (0.058) | | |
| Joint F-test $p$-value | | | | 0.009*** | 0.675 |
| Number of clusters | 70 | 68 | 65 | | |
| **Panel C: Warden Characteristics** | | | | | |
| Warden's age (years) | 34.623 | 35.324 | 34.800 | 0.609 | 0.186 |
| | (0.813) | (0.789) | (0.790) | | |
| Has degree in education | 0.362 | 0.382 | 0.415 | 0.019 | 0.052 |

Continued on next page

*Table A2 continued*

| Variable | (1) Control Mean | (2) T1: Mean | (3) T2: Mean | (4) T1 – Control | (5) T2 – Control |
|---|---|---|---|---|---|
| | (0.058) | (0.059) | (0.062) | | |
| Experience as a warden (years) | 6.029 | 5.757 | 6.138 | -0.357 | 0.177 |
| | (0.655) | (0.592) | (0.588) | | |
| Has received formal training | 0.522 | 0.368 | 0.492 | -0.164* | -0.039 |
| | (0.061) | (0.059) | (0.062) | | |
| Gender attitudes index | -0.012 | -0.066 | -0.146 | -0.054 | -0.129 |
| | (0.122) | (0.108) | (0.119) | | |
| Social norms index | -0.021 | -0.211 | -0.339 | -0.175 | -0.289** |
| | (0.121) | (0.106) | (0.108) | | |
| Joint F-test $p$-value | | | | 0.289 | 0.587 |
| Number of clusters | 69 | 68 | 65 | | |
| **Panel D: KGBV Characteristics** | | | | | |
| Accommodation capacity (students) | 99.938 | 99.933 | 108.000 | 0.685 | 7.995 |
| | (3.236) | (2.274) | (5.047) | | |
| Students currently enrolled | 87.768 | 92.985 | 95.446 | 4.495 | 7.695 |
| | (3.969) | (3.429) | (5.133) | | |
| Distance to Block Development Office | 12.746 | 15.522 | 16.651 | 2.675 | 3.716* |
| | (1.439) | (1.599) | (1.569) | | |
| Number of staff | 6.942 | 7.103 | 6.815 | 0.105 | -0.062 |
| | (0.305) | (0.359) | (0.358) | | |
| Facilities index | -0.057 | -0.002 | 0.036 | 0.040 | 0.127 |
| | (0.118) | (0.126) | (0.126) | | |
| Extracurricular activities index | 0.019 | 0.036 | -0.143 | 0.021 | -0.129 |
| | (0.121) | (0.115) | (0.129) | | |
| Living conditions index | -0.004 | 0.007 | 0.000 | 0.022 | 0.027 |
| | (0.121) | (0.135) | (0.111) | | |
| Joint F-test $p$-value | | | | 0.927 | 0.031** |
| Number of clusters | 69 | 68 | 65 | | |

*Notes:* Student sample includes all baseline observations. Student and household characteristics are averaged within KGBVs. Warden and KGBV characteristics are measured at the KGBV level. Columns (1)–(3) report treatment-arm means; cluster-robust standard errors of the means are reported in parentheses. Columns (4)–(5) report treatment–control differences from regressions that include randomization-stratum fixed effects. Standard errors clustered at the KGBV level. Joint F-test $p$-values test the joint significance of all characteristics reported in the relevant panel. The number-of-clusters row reports the maximum number of KGBVs represented in each panel and treatment arm. Grade 6 is the omitted grade category. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A3: Balance of Baseline Characteristics (Baseline-endline matched Student Sample)

| Variable | (1) Control Mean | (2) T1: Mean | (3) T2: Mean | (4) T1 – Control | (5) T2 – Control |
|---|---|---|---|---|---|
| **Panel A: Student Characteristics and Baseline Outcomes** | | | | | |
| Age (years) | 13.341 | 13.360 | 13.507 | 0.042 | 0.155** |
| | (0.103) | (0.104) | (0.105) | | |
| Share of students in Grade 6 | 0.192 | 0.197 | 0.172 | 0.000 | -0.020 |
| | (0.022) | (0.022) | (0.020) | | |
| Share of students in Grade 7 | 0.208 | 0.195 | 0.195 | -0.018 | -0.011 |
| | (0.022) | (0.021) | (0.025) | | |
| Share of students in Grade 8 | 0.183 | 0.178 | 0.185 | -0.010 | 0.002 |
| | (0.022) | (0.018) | (0.018) | | |
| Share of students in Grade 9 | 0.417 | 0.429 | 0.449 | 0.028 | 0.030 |
| | (0.052) | (0.052) | (0.052) | | |
| Years enrolled in KGBV | 1.870 | 1.745 | 1.915 | -0.123 | 0.045 |
| | (0.085) | (0.078) | (0.068) | | |
| Days absent from school last week | 1.045 | 0.926 | 1.181 | -0.132 | 0.142 |
| | (0.086) | (0.103) | (0.096) | | |
| Daily study hours in hostel | 3.214 | 3.120 | 3.101 | -0.082 | -0.121 |
| | (0.062) | (0.079) | (0.066) | | |
| Total cognitive test score | 11.457 | 12.174 | 11.409 | 0.718 | -0.040 |
| | (0.306) | (0.442) | (0.385) | | |
| English test score | 4.892 | 5.293 | 4.867 | 0.404 | -0.025 |
| | (0.158) | (0.205) | (0.177) | | |
| Previous examination marks (%) | 65.781 | 67.027 | 68.523 | 1.347 | 2.454* |
| | (1.258) | (1.274) | (1.126) | | |
| ASER score | 2.518 | 2.605 | 2.680 | 0.099 | 0.186 |
| | (0.118) | (0.086) | (0.123) | | |
| Perceived barriers index | -0.063 | -0.121 | -0.008 | -0.069 | 0.065 |
| | (0.062) | (0.061) | (0.072) | | |
| Gender attitudes index | 0.008 | 0.116 | 0.056 | 0.118 | 0.047 |
| | (0.056) | (0.072) | (0.054) | | |
| Aspirations index | 0.035 | 0.009 | -0.041 | -0.018 | -0.079 |
| | (0.055) | (0.053) | (0.056) | | |
| Joint F-test $p$-value | | | | 0.338 | 0.201 |
| Number of clusters | 70 | 68 | 64 | | |
| **Panel C: Household Characteristics** | | | | | |
| Number of household members | 6.555 | 6.408 | 6.582 | -0.159 | 0.003 |
| | (0.119) | (0.099) | (0.133) | | |
| Number of siblings | 2.649 | 2.692 | 2.745 | 0.034 | 0.101 |
| | (0.071) | (0.084) | (0.080) | | |
| First-born child | 0.460 | 0.437 | 0.430 | -0.023 | -0.030 |
| | (0.017) | (0.016) | (0.015) | | |
| Father's education (years) | 6.785 | 6.671 | 6.903 | -0.108 | 0.146 |
| | (0.215) | (0.221) | (0.214) | | |
| Mother engaged in farming | 0.620 | 0.682 | 0.593 | 0.062* | -0.023 |
| | (0.022) | (0.023) | (0.027) | | |
| Father engaged in farming | 0.525 | 0.577 | 0.501 | 0.052 | -0.022 |
| | (0.027) | (0.028) | (0.026) | | |
| Number of rooms in household | 2.595 | 2.548 | 2.589 | -0.053 | -0.006 |
| | (0.077) | (0.056) | (0.064) | | |
| Joint F-test $p$-value | | | | 0.114 | 0.783 |
| Number of clusters | 70 | 68 | 64 | | |
| **Panel C: Warden Characteristics** | | | | | |
| Warden's age (years) | 34.623 | 35.324 | 34.800 | 0.609 | 0.186 |
| | (0.813) | (0.789) | (0.790) | | |
| Has degree in education | 0.362 | 0.382 | 0.415 | 0.019 | 0.052 |

Continued on next page

*Table A3 continued*

| Variable | (1) Control Mean | (2) T1: Mean | (3) T2: Mean | (4) T1 – Control | (5) T2 – Control |
|---|---|---|---|---|---|
| | (0.058) | (0.059) | (0.062) | | |
| Experience as a warden (years) | 6.029 | 5.757 | 6.138 | -0.357 | 0.177 |
| | (0.655) | (0.592) | (0.588) | | |
| Has received formal training | 0.522 | 0.368 | 0.492 | -0.164* | -0.039 |
| | (0.061) | (0.059) | (0.062) | | |
| Gender attitudes index | -0.012 | -0.066 | -0.146 | -0.054 | -0.129 |
| | (0.122) | (0.108) | (0.119) | | |
| Social norms index | -0.021 | -0.211 | -0.339 | -0.175 | -0.289** |
| | (0.121) | (0.106) | (0.108) | | |
| Joint F-test $p$-value | | | | 0.289 | 0.587 |
| Number of clusters | 69 | 68 | 65 | | |
| **Panel D: KGBV Characteristics** | | | | | |
| Hostel accommodation capacity | 99.938 | 99.933 | 108.000 | 0.685 | 7.995 |
| | (3.236) | (2.274) | (5.047) | | |
| Students currently enrolled | 87.768 | 92.985 | 95.446 | 4.495 | 7.695 |
| | (3.969) | (3.429) | (5.133) | | |
| Distance to Block Development Office | 12.746 | 15.522 | 16.651 | 2.675 | 3.716* |
| | (1.439) | (1.599) | (1.569) | | |
| Number of staff | 6.942 | 7.103 | 6.815 | 0.105 | -0.062 |
| | (0.305) | (0.359) | (0.358) | | |
| Facilities index | -0.057 | -0.002 | 0.036 | 0.040 | 0.127 |
| | (0.118) | (0.126) | (0.126) | | |
| Extracurricular activities index | 0.019 | 0.036 | -0.143 | 0.021 | -0.129 |
| | (0.121) | (0.115) | (0.129) | | |
| Living conditions index | -0.004 | 0.007 | 0.000 | 0.022 | 0.027 |
| | (0.121) | (0.135) | (0.111) | | |
| Joint F-test $p$-value | | | | 0.927 | 0.031** |
| Number of clusters | 69 | 68 | 65 | | |

*Notes:* Student sample includes baseline observations tracked at endline. Student and household characteristics are averaged within KGBVs. Warden and KGBV characteristics are measured at the KGBV level. Columns (1)–(3) report treatment-arm means; cluster-robust standard errors of the means are reported in parentheses. Columns (4)–(5) report treatment–control differences from regressions that include randomization-stratum fixed effects. Standard errors clustered at the KGBV level. Joint F-test $p$-values test the joint significance of all characteristics reported in the relevant panel. The number-of-clusters row reports the maximum number of KGBVs represented in each panel and treatment arm. Grade 6 is the omitted grade category. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A4: Workshop and Coaching Call Attendance

| Variable | KGBVs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| Workshop 1 attendance days | 133 | 3.32 | 1.37 | 0 | 4 |
| Workshop 2 attendance days | 133 | 3.61 | 1.00 | 0 | 4 |
| Workshop 3 attendance days | 133 | 2.07 | 1.29 | 0 | 3 |
| Coaching Call Cycle 1 attendance | 133 | 9.02 | 3.68 | 0 | 13 |
| Coaching Call Cycle 2 attendance | 133 | 8.89 | 2.72 | 0 | 12 |
| Total coaching call attendance | 133 | 17.91 | 5.32 | 0 | 25 |

*Notes:* Attendance is measured at the KGBV level and captures participation by any warden from the KGBV during the intervention period. Workshops 1 and 2 were four-day training workshops, while Workshop 3 was a three-day training workshop. Coaching Call Cycles 1 and 2 comprised a maximum of 13 and 12 calls, respectively, for a total maximum of 25 coaching calls.

Table A5: Descriptive Statistics on Lesson Exposure, Participation, and Comprehension

| **Variable** | **Mean** | **SD** | **Min** | **Max** |
|---|---|---|---|---|
| *Panel A: Session Attendance* | | | | |
| English sessions attended (midpoint of range; 0–46) | 31 | 14.454 | 0.0 | 42 |
| Life-skills sessions attended (midpoint of range; 0–23) | 15 | 7.108 | 0.0 | 21 |
| *Panel B: Modules Taught* | | | | |
| Self-awareness | 0.49 | 0.50 | 0.0 | 1.0 |
| Health, hygiene, and safety | 0.52 | 0.50 | 0.0 | 1.0 |
| Life skills | 0.54 | 0.50 | 0.0 | 1.0 |
| Gender sensitivity | 0.46 | 0.50 | 0.0 | 1.0 |
| Social awareness | 0.41 | 0.49 | 0.0 | 1.0 |
| *Panel C: Active Participation* | | | | |
| Self-awareness | 0.47 | 0.50 | 0.0 | 1.0 |
| Health, hygiene, and safety | 0.48 | 0.50 | 0.0 | 1.0 |
| Life skills | 0.53 | 0.50 | 0.0 | 1.0 |
| Gender sensitivity | 0.42 | 0.49 | 0.0 | 1.0 |
| Social awareness | 0.39 | 0.49 | 0.0 | 1.0 |
| *Panel D: Understanding of Lesson Content* | | | | |
| Health, hygiene, and safety | 0.91 | 0.28 | 0.0 | 1.0 |
| Self-awareness | 0.91 | 0.29 | 0.0 | 1.0 |
| Life skills | 0.91 | 0.28 | 0.0 | 1.0 |
| Gender sensitivity | 0.88 | 0.32 | 0.0 | 1.0 |
| Social awareness | 0.89 | 0.31 | 0.0 | 1.0 |

*Notes:* The table reports descriptive statistics for students in the treatment arms. Session-attendance measures use the midpoint of the reported attendance range. Variables in Panels B–D are binary indicators. Panel B indicates whether the relevant module was taught, Panel C indicates whether the student actively participated in the module, and Panel D indicates whether the student reported understanding the lesson content.

Table A6: Classroom Observation Measures

| Variable | KGBVs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Round 1** | | | | | |
| **A. Time on Task** | | | | | |
| Percentage of Time Warden on Task | 124 | 83.74 | 26.84 | 0 | 100 |
| Percentage of Time Student on Task | 124 | 78.90 | 27.23 | 0 | 100 |
| **B. Quality of Teaching Practices** | | | | | |
| Supportive Classroom Environment | 124 | 6.88 | 2.41 | 0 | 10 |
| Positive Classroom Behaviour | 124 | 6.64 | 2.49 | 0 | 10 |
| Lesson Facilitation | 124 | 6.77 | 2.36 | 0 | 10 |
| Checks for Understanding | 124 | 6.44 | 2.65 | 0 | 10 |
| Feedback | 124 | 6.00 | 2.58 | 0 | 10 |
| Critical Thinking | 124 | 6.13 | 2.43 | 0 | 10 |
| Autonomy | 124 | 6.46 | 2.66 | 0 | 10 |
| Perseverance | 124 | 6.28 | 2.48 | 0 | 10 |
| Social and Collaborative Skills | 124 | 5.06 | 3.02 | 0 | 10 |
| **Round 2** | | | | | |
| **A. Time on Task** | | | | | |
| Percentage of Time Warden on Task | 98 | 77.38 | 33.78 | 0 | 100 |
| Percentage of Time Student on Task | 98 | 74.49 | 35.83 | 0 | 100 |
| **B. Quality of Teaching Practices** | | | | | |
| Supportive Classroom Environment | 98 | 6.91 | 2.43 | 0 | 10 |
| Positive Classroom Behaviour | 98 | 6.64 | 2.61 | 0 | 10 |
| Lesson Facilitation | 98 | 6.47 | 2.66 | 0 | 10 |
| Checks for Understanding | 98 | 6.39 | 2.91 | 0 | 10 |
| Feedback | 98 | 6.15 | 2.77 | 0 | 10 |
| Critical Thinking | 98 | 6.02 | 2.94 | 0 | 10 |
| Autonomy | 98 | 6.28 | 2.80 | 0 | 10 |
| Perseverance | 98 | 5.98 | 2.83 | 0 | 10 |
| Social and Collaborative Skills | 98 | 5.53 | 3.13 | 0 | 10 |

*Notes:* The classroom observation tool was administered in two rounds. Time-on-task measures capture the percentage of observed class time during which the warden and students, respectively, were engaged in the intended lesson activity and range from 0 to 100. Quality-of-teaching-practice measures are scored on a scale from 0 to 10, with higher scores indicating stronger observed teaching practices. The smaller number of observations in Round 2 primarily reflects scheduled sessions that were canceled or did not take place on the day of the observation visit.

Table A7: Weekly Parent IVR Implementation and Engagement

| Week | Parents Targeted | Calls Answered | Answer Rate (%) | Listened ≥50% | Successful Engagement (%) |
|---|---|---|---|---|---|
| Week 1 | 952 | 766 | 80.46 | 382 | 40.13 |
| Week 2 | 953 | 756 | 79.33 | 298 | 31.27 |
| Week 3 | 3441 | 2850 | 82.82 | 1592 | 46.27 |
| Week 4 | 2443 | 2043 | 83.63 | 1043 | 42.69 |
| Week 5 | 3441 | 2792 | 81.14 | 1450 | 42.14 |
| Week 6 | 3441 | 2780 | 80.79 | 1244 | 36.15 |
| Week 7 | 3634 | 2947 | 81.10 | 1281 | 35.25 |
| Week 8 | 3634 | 2925 | 80.49 | 1279 | 35.20 |
| Week 9 | 4277 | 3517 | 82.23 | 1219 | 28.50 |
| Week 10 | 4277 | 3471 | 81.16 | 1368 | 31.99 |
| Week 11 | 4277 | 3258 | 76.17 | 1106 | 25.86 |
| Week 12 | 4277 | 3611 | 84.43 | 1650 | 38.58 |
| Week 13 | 4277 | 3563 | 83.31 | 1415 | 33.08 |
| Week 14 | 4277 | 3404 | 79.59 | 1268 | 29.65 |
| Week 15 | 4278 | 3249 | 75.95 | 1278 | 29.87 |
| Week 16 | 4278 | 3201 | 74.82 | 1686 | 39.41 |
| Week 17 | 4278 | 2997 | 70.06 | 1267 | 29.62 |
| Week 18 | 4278 | 3299 | 77.12 | 1852 | 43.29 |
| Week 19 | 4278 | 3319 | 77.58 | 1709 | 39.95 |
| Week 20 | 4278 | 3236 | 75.64 | 1708 | 39.93 |
| Week 21 | 4278 | 3146 | 73.54 | 1746 | 40.81 |
| Week 22 | 4278 | 3166 | 74.01 | 1540 | 36.00 |
| Week 23 | 4278 | 3129 | 73.14 | 1458 | 34.08 |
| Week 24 | 4278 | 3131 | 73.19 | 1327 | 31.02 |
| Week 25 | 4276 | 3100 | 72.50 | 1291 | 30.19 |
| Week 26 | 4276 | 3044 | 71.19 | 1164 | 27.22 |
| Week 27 | 4276 | 3075 | 71.91 | 1356 | 31.71 |
| Week 28 | 4276 | 2759 | 64.52 | 987 | 23.08 |
| Week 29 | 4276 | 2753 | 64.38 | 1167 | 27.29 |
| Week 30 | 4276 | 2775 | 64.90 | 1162 | 27.17 |
| Week 31 | 4276 | 2814 | 65.81 | 961 | 22.47 |
| Week 32 | 4276 | 2835 | 66.30 | 1154 | 26.99 |
| Week 33 | 4276 | 2837 | 66.35 | 1097 | 25.65 |
| Week 34 | 4276 | 3079 | 72.01 | 1524 | 35.64 |
| Week 35 | 4276 | 3080 | 72.03 | 1548 | 36.20 |
| Week 36 | 4276 | 3024 | 70.72 | 1408 | 32.93 |
| Week 37 | 4276 | 3004 | 70.25 | 1289 | 30.14 |
| Week 38 | 4276 | 3060 | 71.56 | 1310 | 30.64 |

*Continued on next page*

*Table A7 continued from previous page*

| Week | Parents Targeted | Calls Answered | Answer Rate (%) | Listened ≥50% | Successful Engagement (%) |
|---|---|---|---|---|---|
| Week 39 | 4276 | 2970 | 69.46 | 1313 | 30.71 |
| Week 40 | 4276 | 2945 | 68.87 | 1326 | 31.01 |
| Week 41 | 4276 | 2908 | 68.01 | 1301 | 30.43 |
| Week 42 | 4276 | 2905 | 67.94 | 1168 | 27.32 |
| Week 43 | 4276 | 2942 | 68.80 | 1189 | 27.81 |
| Week 44 | 4243 | 3373 | 79.50 | 1582 | 37.28 |
| Week 45 | 4243 | 3313 | 78.08 | 1476 | 34.79 |
| Week 46 | 4243 | 3364 | 79.28 | 1293 | 30.47 |
| Week 47 | 4243 | 3271 | 77.09 | 1371 | 32.31 |
| Week 48 | 4243 | 3183 | 75.02 | 1191 | 28.07 |
| Week 49 | 4243 | 3113 | 73.37 | 1213 | 28.59 |
| Week 50 | 4243 | 3067 | 72.28 | 1107 | 26.09 |
| Week 51 | 4243 | 3043 | 71.72 | 937 | 22.08 |
| Week 52 | 4243 | 3033 | 71.48 | 990 | 23.33 |

*Notes:* Calls are classified as answered when the recorded listening duration is greater than zero. Successful engagement is defined as listening to at least 50% of the weekly IVR audio message. Both rates are calculated relative to the number of parent phone numbers targeted in the corresponding week.

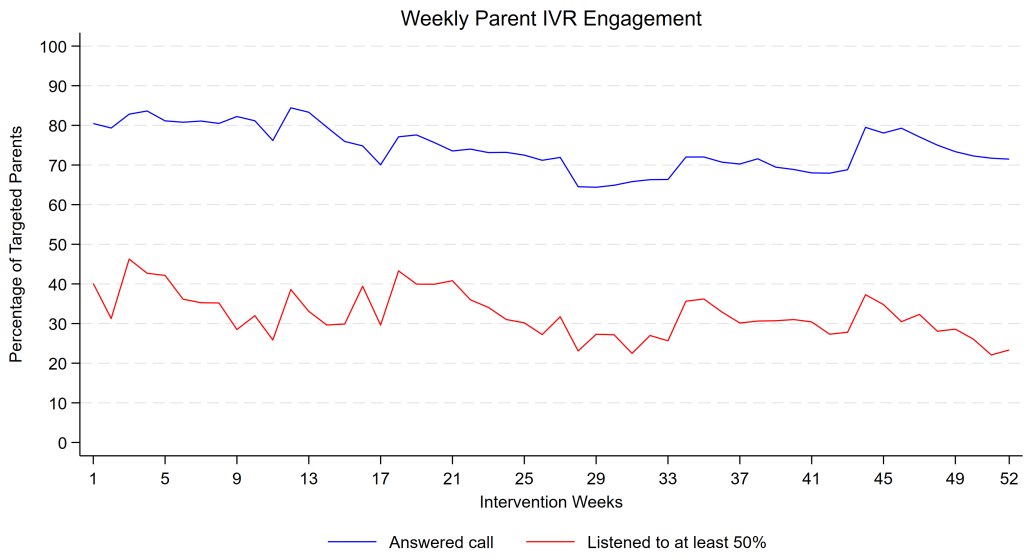



Figure 6: Weekly Parent IVR Reach and Engagement

Table A8: Parent-Reported Engagement with IVR audio messages at Endline

| Variable | Obs. | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| **Panel A: IVR Reach and Engagement** | | | | | |
| Household member received an IVR call | 1,384 | 0.564 | 0.496 | 0 | 1 |
| Household member saved the IVR number | 1,384 | 0.287 | 0.452 | 0 | 1 |
| IVR compliance | 1,384 | 0.574 | 0.495 | 0 | 1 |
| Weekly listening frequency | 794 | 0.388 | 0.488 | 0 | 1 |
| Used the call-back option | 794 | 0.267 | 0.443 | 0 | 1 |
| Number of successful IVR weeks ($\geq$50%) | 1,163 | 15.120 | 12.977 | 0 | 51 |
| **Panel B: Features Liked** | | | | | |
| Liked the language | 794 | 0.661 | 0.474 | 0 | 1 |
| Liked the ideas/content | 794 | 0.505 | 0.500 | 0 | 1 |
| Liked the examples | 794 | 0.335 | 0.472 | 0 | 1 |
| Liked the tone | 794 | 0.253 | 0.435 | 0 | 1 |
| None of the listed features | 794 | 0.189 | 0.392 | 0 | 1 |
| **Panel C: Features Disliked** | | | | | |
| Disliked the length | 794 | 0.280 | 0.449 | 0 | 1 |
| Disliked repetition | 794 | 0.253 | 0.435 | 0 | 1 |
| Disliked relevance | 794 | 0.164 | 0.370 | 0 | 1 |
| Disliked audio quality | 794 | 0.126 | 0.332 | 0 | 1 |
| None of the listed features | 794 | 0.446 | 0.497 | 0 | 1 |
| **Panel D: Perceived Usefulness** | | | | | |
| Found the IVR calls helpful | 794 | 0.676 | 0.468 | 0 | 1 |

*Notes:* Except for the number of successful IVR weeks, all measures are binary indicators coded from 0 to 1; for these measures, the mean represents the proportion of respondents for whom the indicator equals one. IVR compliance is coded as one if either a household member reported receiving an IVR call or a household member had saved the IVR number. The number of successful IVR weeks is based on matched administrative IVR records and measures the number of weeks in which the parent listened to at least 50% of the audio message. Questions on features liked and disliked allowed multiple responses. The number of observations varies across measures because of survey skip patterns and the availability of matched administrative IVR records.

Table A9: Effects on Cognitive Outcomes

| | (1) English Score Grades 9–10 | (2) English Score Grades 7–8 | (3) English Score Pooled | (4) Final Marks | (5) Raven Matrices Score |
|---|---|---|---|---|---|
| Any treatment | 0.161** | 0.026 | 0.092 | 0.069 | 0.052 |
| | (0.080) | (0.122) | (0.080) | (0.057) | (0.065) |
| | [0.045] | [0.833] | [0.252] | [0.230] | [0.426] |
| T1 (Learning Program) | 0.120 | 0.062 | 0.091 | 0.013 | -0.021 |
| | (0.096) | (0.129) | (0.087) | (0.068) | (0.073) |
| | [0.210] | [0.631] | [0.299] | [0.849] | [0.777] |
| T2 (Learning Program + Parental IVR) | 0.203** | -0.012 | 0.093 | 0.126** | 0.127 |
| | (0.091) | (0.139) | (0.091) | (0.061) | (0.081) |
| | [0.027] | [0.932] | [0.306] | [0.042] | [0.117] |
| p-value: T1 = T2 | 0.404 | 0.511 | 0.974 | 0.079 | 0.080 |
| Observations | 4621 | 4621 | 9243 | 8730 | 9243 |
| Control Mean | 0.000 | -0.000 | 0.000 | -0.000 | -0.000 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes |
| Clusters | 198 | 134 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' cognitive outcomes. The first specification pools the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Outcomes are standardized relative to the control group. All regressions include controls and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A10: Effects on Confidence

| | (1) Confidence Index | (2) Leadership Confidence | (3) Public Confidence | (4) Private Confidence |
|---|---|---|---|---|
| Any treatment | 0.130** | 4.262* | 1.373 | 1.487 |
| | (0.061) | (2.526) | (2.482) | (1.549) |
| | [0.033] | [0.093] | [0.581] | [0.338] |
| T1 (Learning Program) | 0.145** | 2.862 | 3.349 | 1.604 |
| | (0.072) | (2.991) | (2.901) | (1.637) |
| | [0.046] | [0.340] | [0.250] | [0.328] |
| T2 (Learning Program + Parental IVR) | 0.116 | 5.715* | -0.684 | 1.373 |
| | (0.071) | (2.988) | (2.921) | (2.016) |
| | [0.104] | [0.057] | [0.815] | [0.496] |
| p-value: T1 = T2 | 0.705 | 0.373 | 0.185 | 0.908 |
| Observations | 6867 | 8678 | 8712 | 7510 |
| Control Mean | 0.000 | 69.544 | 32.729 | 73.109 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 199 | 202 | 202 | 200 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' confidence. Column (1) reports effects on the standardized confidence index; Columns (2)–(4) report effects on its components in their original units. The pooled specification combines the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Regressions include the same basic controls used for cognitive outcomes along with a social desirability bias index and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A11: Effects on Aspirations

| | (1) Aspirations Index | (2) Study Beyond Grade 12 | (3) High-Level Occupation | (4) Marriage After Age 25 |
|---|---|---|---|---|
| Any treatment | 0.085* | 0.045*** | 0.021 | 0.003 |
| | (0.047) | (0.016) | (0.013) | (0.022) |
| | [0.071] | [0.004] | [0.113] | [0.898] |
| T1 (Learning Program) | 0.053 | 0.038** | 0.023 | -0.017 |
| | (0.052) | (0.019) | (0.015) | (0.025) |
| | [0.313] | [0.042] | [0.125] | [0.511] |
| T2 (Learning Program + Parental IVR) | 0.119** | 0.053*** | 0.018 | 0.023 |
| | (0.058) | (0.019) | (0.016) | (0.026) |
| | [0.043] | [0.006] | [0.258] | [0.380] |
| $p$-value: T1 = T2 | 0.268 | 0.497 | 0.731 | 0.158 |
| Observations | 9,243 | 8,947 | 8,872 | 8,955 |
| Control Mean | 0.000 | 0.770 | 0.722 | 0.764 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' aspirations. Column (1) reports effects on overall aspirations index, Columns (2)–(4) report effects on its individual components. The pooled specification combines the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Regressions include the same basic controls used for cognitive outcomes along with a social desirability bias index and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A12: Effects on Gender Attitudes

| | (1) Gender Attitude Index | (2) Education Attitude Index | (3) Equal Rights Attitude Index | (4) Employment Attitude Index |
|---|---|---|---|---|
| Any treatment | 0.070 | 0.105** | 0.029 | 0.080 |
| | (0.086) | (0.043) | (0.047) | (0.059) |
| | [0.418] | [0.016] | [0.541] | [0.178] |
| T1 (Learning Program) | -0.011 | 0.123** | 0.084* | 0.050 |
| | (0.093) | (0.050) | (0.049) | (0.068) |
| | [0.909] | [0.016] | [0.090] | [0.462] |
| T2 (Learning Program + Parental IVR) | 0.154 | 0.087* | -0.029 | 0.112 |
| | (0.102) | (0.050) | (0.064) | (0.068) |
| | [0.132] | [0.085] | [0.652] | [0.101] |
| $p$-value: T1 = T2 | 0.091 | 0.490 | 0.085 | 0.356 |
| Observations | 9,243 | 9,243 | 9,243 | 9,243 |
| Control Mean | -0.000 | -0.000 | 0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' gender attitudes. Column (1) reports effects on the overall gender attitude index, Columns (2)–(4) report effects on the domain-specific indices relating to girls' education, equal rights, and employment. The pooled specification combines the two treatment arms, and the subsequent specifications report treatment-arm-specific effects. Outcomes are standardized relative to the control group. Regressions include the same basic controls used for cognitive outcomes along with a social desirability bias index and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A13: Effects on Social and Emotional Learning (SEL) Outcomes

| | (1) SEL Index | (2) Social Awareness Index | (3) Self-Awareness Index | (4) Self-Management Index |
|---|---|---|---|---|
| Any treatment | 0.056 | 0.091 | -0.018 | 0.040 |
| | (0.068) | (0.061) | (0.058) | (0.064) |
| | [0.413] | [0.132] | [0.758] | [0.532] |
| T1 (Learning Program) | 0.004 | 0.049 | -0.042 | -0.015 |
| | (0.075) | (0.067) | (0.061) | (0.072) |
| | [0.956] | [0.468] | [0.496] | [0.841] |
| T2 (Learning Program + Parental IVR) | 0.109 | 0.136* | 0.007 | 0.097 |
| | (0.083) | (0.074) | (0.071) | (0.079) |
| | [0.189] | [0.068] | [0.927] | [0.220] |
| *p-value*: T1 = T2 | 0.203 | 0.237 | 0.447 | 0.175 |
| Observations | 9243 | 9243 | 9243 | 9243 |
| Control Mean | 0.000 | 0.000 | 0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' social and emotional learning outcomes. Column (1) reports effects on the overall SEL index, while Columns (2)–(4) report effects on the self-awareness, social-awareness, and self-management sub-indices. The pooled specification combines the two treatment arms; the subsequent specifications report treatment-arm-specific effects. All outcomes are standardized relative to the control group. Regressions include the same basic controls used for cognitive outcomes along with a social desirability bias index and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A14: Effects on Other Secondary Outcomes

| | (1) Headache or Stomach-Ache | (2) Feeling Stressed | (3) Trouble Concentrating | (4) Anxiety Index | (5) Growth Mindset Index | (6) Social Norms Index |
|---|---|---|---|---|---|---|
| Any treatment | 0.018 | -0.022 | -0.002 | 0.073 | 0.020 | 0.040 |
| | (0.025) | (0.025) | (0.026) | (0.065) | (0.057) | (0.055) |
| | [0.471] | [0.375] | [0.947] | [0.269] | [0.719] | [0.469] |
| T1 (Learning Program) | 0.026 | -0.016 | 0.029 | 0.046 | -0.037 | 0.094 |
| | (0.026) | (0.027) | (0.027) | (0.068) | (0.065) | (0.067) |
| | [0.303] | [0.557] | [0.282] | [0.503] | [0.569] | [0.162] |
| T2 (Learning Program + Parental IVR) | 0.009 | -0.028 | -0.034 | 0.101 | 0.080 | -0.016 |
| | (0.032) | (0.032) | (0.030) | (0.084) | (0.068) | (0.069) |
| | [0.772] | [0.373] | [0.267] | [0.234] | [0.237] | [0.815] |
| p-value: T1 = T2 | 0.564 | 0.706 | 0.028 | 0.490 | 0.100 | 0.172 |
| Observations | 9056 | 8976 | 8890 | 9243 | 9243 | 9243 |
| Control Mean | 0.470 | 0.523 | 0.496 | 0.000 | -0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on other secondary student outcomes. Columns (1)–(3) report indicators for experiencing headaches or stomach-aches, feeling stressed, and having trouble concentrating. Columns (4)–(6) report standardized indices of anxiety, growth mindset, and perceived social norms. The pooled specification combines the two treatment arms; the subsequent specifications report treatment-arm-specific effects. Regressions include basic controls and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A15: Effects on Knowledge of Key Life-Skills Content

| | (1) Gender Role | (2) Nontraditional Gender Identity | (3) Role Model | (4) Empathy | (5) Goal Setting | (6) Menstruation Myth | (7) Bad Touch |
|---|---|---|---|---|---|---|---|
| Any treatment | 0.127 | 0.025 | 0.029 | 0.082*** | 0.055* | 0.052 | 0.011 |
| | (0.108) | (0.021) | (0.018) | (0.028) | (0.030) | (0.036) | (0.031) |
| | [0.244] | [0.233] | [0.120] | [0.004] | [0.074] | [0.150] | [0.731] |
| T1 (Learning Program) | 0.023 | 0.009 | 0.025 | 0.060* | 0.035 | 0.035 | -0.015 |
| | (0.114) | (0.021) | (0.019) | (0.031) | (0.034) | (0.038) | (0.034) |
| | [0.839] | [0.657] | [0.195] | [0.053] | [0.306] | [0.367] | [0.669] |
| T2 (Learning Program + Parental IVR) | 0.234* | 0.041 | 0.032 | 0.105*** | 0.075* | 0.069 | 0.037 |
| | (0.136) | (0.026) | (0.025) | (0.031) | (0.038) | (0.044) | (0.036) |
| | [0.086] | [0.124] | [0.192] | [0.001] | [0.053] | [0.121] | [0.302] |
| $p$-value: T1 = T2 | 0.105 | 0.200 | 0.771 | 0.099 | 0.334 | 0.417 | 0.118 |
| Observations | 9,070 | 9,176 | 8,736 | 8,910 | 8,869 | 8,663 | 8,841 |
| Control Mean | -0.000 | 0.789 | 0.237 | 0.618 | 0.349 | 0.335 | 0.603 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 202 | 202 | 202 | 202 | 202 | 202 | 202 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on students' knowledge of key life-skills content. Column (1) reports the standardized gender-role knowledge measure, while Columns (2)–(7) report outcomes in their original units. The pooled specification combines the two treatment arms; the subsequent specifications report treatment-arm-specific effects. All regressions include the same basic controls and stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A16: Cognitive Outcomes by English-Session Attendance

| | (1) English Test Score Grades 9–10 | (2) English Test Score Grades 7–8 | (3) English Test Score Grades 7–10 | (4) Final Marks | (5) Raven Matrices Score |
|---|---|---|---|---|---|
| T1 × No Sessions | -0.077 | -0.019 | -0.032 | -0.106 | -0.230* |
| | (0.165) | (0.180) | (0.124) | (0.108) | (0.124) |
| T1 × 1–27 Sessions | 0.085 | -0.271* | -0.089 | -0.165 | -0.110 |
| | (0.105) | (0.159) | (0.094) | (0.104) | (0.081) |
| T1 × 28–46 Sessions | 0.195* | 0.191 | 0.205** | 0.137** | 0.066 |
| | (0.105) | (0.116) | (0.081) | (0.064) | (0.074) |
| T2 × No Sessions | -0.008 | -0.376* | -0.179 | 0.093 | -0.179* |
| | (0.138) | (0.192) | (0.126) | (0.089) | (0.102) |
| T2 × 1–27 Sessions | 0.082 | -0.219** | -0.061 | 0.017 | -0.002 |
| | (0.114) | (0.110) | (0.085) | (0.069) | (0.088) |
| T2 × 28–46 Sessions | 0.295*** | 0.100 | 0.207** | 0.190*** | 0.211*** |
| | (0.088) | (0.135) | (0.086) | (0.062) | (0.080) |
| Observations | 4,621 | 4,621 | 9,243 | 8,730 | 9,243 |
| Control Mean | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes |
| Clusters | 198 | 134 | 202 | 202 | 202 |

*Notes:* The table reports estimates of cognitive outcomes by treatment arm and English-session attendance. Attendance is divided into three categories: no sessions, 1–27 sessions, and 28–46 sessions. The omitted category is the control group. All outcomes are standardized relative to the control group. All regressions include basic controls and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A17: Non-Cognitive Outcomes by Life-Skills-Session Attendance

| | (1) Confidence Index | (2) Aspirations Index | (3) Gender Attitudes Index | (4) SEL Index |
|---|---|---|---|---|
| T1 × No Sessions | -0.011 | -0.180 | -0.030 | -0.310*** |
| | (0.134) | (0.113) | (0.135) | (0.113) |
| T1 × 1–13 Sessions | -0.012 | -0.006 | -0.057 | -0.012 |
| | (0.090) | (0.062) | (0.083) | (0.070) |
| T1 × 14–23 Sessions | 0.167** | 0.134*** | 0.039 | 0.059 |
| | (0.071) | (0.052) | (0.094) | (0.073) |
| T2 × No Sessions | -0.327 | -0.137* | -0.171 | -0.435*** |
| | (0.213) | (0.077) | (0.186) | (0.109) |
| T2 × 1–13 Sessions | -0.020 | -0.029 | -0.081 | -0.106 |
| | (0.081) | (0.065) | (0.079) | (0.078) |
| T2 × 14–23 Sessions | 0.180*** | 0.232*** | 0.299*** | 0.236*** |
| | (0.066) | (0.051) | (0.098) | (0.079) |
| Observations | 6,867 | 9,243 | 9,243 | 9,243 |
| Control Mean | 0.000 | 0.000 | 0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 199 | 202 | 202 | 202 |

*Notes:* The table reports estimates for non-cognitive primary outcomes by treatment arm and the number of life-skills sessions attended. Session attendance is divided into three categories: no sessions, 1–13 sessions, and 14–23 sessions. All outcomes are standardized. Regressions include basic controls and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A18: Instrumental-Variables Estimates for Cognitive Outcomes

| | (1) English Test Score Grades 9–10 | (2) English Test Score Grades 7–8 | (3) Final Marks | (4) Raven Matrices Score |
|---|---|---|---|---|
| English Sessions Attended (Pooled) | 0.006** | 0.001 | 0.003 | 0.002 |
| | (0.003) | (0.004) | (0.002) | (0.002) |
| | [0.017] | [0.829] | [0.156] | [0.354] |
| English Sessions Attended in T1 | 0.005 | 0.002 | 0.001 | 0.000 |
| | (0.003) | (0.004) | (0.002) | (0.002) |
| | [0.130] | [0.616] | [0.603] | [0.971] |
| English Sessions Attended in T2 | 0.008** | -0.001 | 0.004** | 0.004 |
| | (0.003) | (0.004) | (0.002) | (0.003) |
| | [0.011] | [0.905] | [0.036] | [0.114] |
| $p$-value: T1 = T2 | 0.474 | 0.467 | 0.175 | 0.129 |
| First-stage F-statistic | 2,227.9 | | | |
| Observations | 4,241 | 4,373 | 8,148 | 8,615 |
| Control Mean | 0.000 | 0.000 | 0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 187 | 134 | 202 | 202 |

*Notes:* The table reports two-stage least squares estimates of the effect of English-session attendance on students' cognitive outcomes. In the pooled specification, English-session attendance is instrumented by assignment to either treatment arm. In the arm-specific specification, English-session attendance in T1 and T2 is instrumented by assignment to the corresponding treatment arms. Coefficients represent the estimated effect of attending one extra English session. Outcomes are standardized relative to the control group. Regressions include basic controls and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. The reported $p$-value tests equality of the T1 and T2 coefficients. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A19: Instrumental-Variables Estimates for Non-Cognitive Outcomes

| | (1) Confidence Index | (2) Aspirations Index | (3) Gender Attitudes Index | (4) SEL Index |
|---|---|---|---|---|
| Life-Skills Sessions Attended (Pooled) | 0.008** | 0.007** | 0.006 | 0.004 |
| | (0.004) | (0.003) | (0.006) | (0.004) |
| | [0.034] | [0.024] | [0.310] | [0.322] |
| Life-Skills Sessions Attended in T1 | 0.009* | 0.005 | 0.000 | 0.001 |
| | (0.005) | (0.003) | (0.006) | (0.005) |
| | [0.072] | [0.188] | [0.936] | [0.759] |
| Life-Skills Sessions Attended in T2 | 0.008* | 0.009** | 0.011* | 0.007 |
| | (0.005) | (0.003) | (0.006) | (0.005) |
| | [0.079] | [0.012] | [0.090] | [0.175] |
| $p$-value: T1 = T2 | 0.901 | 0.264 | 0.098 | 0.294 |
| First-stage F-statistic | 1,656.0 | | | |
| Observations | 6,435 | 8,619 | 8,619 | 8,619 |
| Control Mean | 0.000 | 0.000 | 0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 198 | 202 | 202 | 202 |

*Notes:* The table reports two-stage least squares estimates of the effect of life-skills-session attendance on students' non-cognitive outcomes. In the pooled specification, life-skills-session attendance is instrumented by assignment to either treatment arm. In the arm-specific specification, life-skills-session attendance in T1 and T2 is instrumented by assignment to the corresponding treatment arms. Coefficients represent the estimated effect of attending one additional life-skills session. All outcomes are standardized relative to the control group. All regressions include basic controls along with a social desirability bias index and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional $p$-values are reported in square brackets. The reported $p$-value tests equality of the T1 and T2 coefficients. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A20: Effects on Primary Outcomes by Household Education

| | (1) English Score Grades 9–10 | (2) English Score Grades 7–8 | (3) Final Marks | (4) Confidence Index | (5) Aspirations Index | (6) Gender Attitudes Index | (7) SEL Index |
|---|---|---|---|---|---|---|---|
| T1 × Below Grade 10 | 0.134 | 0.056 | -0.006 | 0.225*** | 0.048 | -0.049 | -0.077 |
| | (0.108) | (0.144) | (0.072) | (0.081) | (0.071) | (0.125) | (0.088) |
| T1 × Grade 10 or Above | 0.042 | -0.003 | -0.013 | 0.020 | -0.014 | -0.018 | -0.005 |
| | (0.097) | (0.124) | (0.077) | (0.080) | (0.060) | (0.086) | (0.078) |
| T2 × Below Grade 10 | 0.225** | 0.075 | 0.162** | 0.143* | 0.124* | 0.248** | 0.100 |
| | (0.101) | (0.151) | (0.068) | (0.076) | (0.073) | (0.126) | (0.091) |
| T2 × Grade 10 or Above | 0.127 | -0.087 | 0.074 | 0.020 | 0.043 | 0.044 | 0.055 |
| | (0.101) | (0.136) | (0.064) | (0.083) | (0.061) | (0.096) | (0.085) |
| Observations | 3,484 | 3,499 | 6,625 | 5,241 | 6,983 | 6,983 | 6,983 |
| Control Mean | 0.050 | 0.037 | 0.042 | 0.040 | 0.063 | 0.046 | 0.047 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 196 | 130 | 201 | 198 | 202 | 202 | 202 |

*Notes:* The table reports treatment effects by household education. A household is classified as having education at Grade 10 or above when either parent completed at least Grade 10; otherwise, it is classified as below Grade 10. The regressions include the household-education main effect, so each reported coefficient represents the treatment effect within the indicated household-education subgroup. Outcomes are standardized. All specifications include basic controls and strata fixed effects. Column (3) additionally includes grade fixed effects, and Columns (4)–(7) control for a social desirability bias index. Standard errors clustered at the KGBV level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A21: Robustness: Effects on Primary Outcomes (Endline Student)

| | **Cognitive Outcomes** | | | **Non-cognitive Outcomes** | | | |
|---|---|---|---|---|---|---|---|
| | (1) English Score (Grades 9–10) | (2) English Score (Grades 7–8) | (3) Final Marks | (4) Confidence Index | (5) Aspirations Index | (6) Gender Attitudes Index | (7) SEL Index |
| Any treatment | 0.186* | 0.063 | 0.063 | 0.128* | 0.056 | 0.013 | 0.085 |
| | (0.099) | (0.105) | (0.052) | (0.065) | (0.047) | (0.081) | (0.070) |
| | [0.062] | [0.548] | [0.227] | [0.052] | [0.230] | [0.868] | [0.223] |
| T1 (Learning Program) | 0.151 | 0.112 | 0.024 | 0.094 | 0.030 | -0.070 | 0.054 |
| | (0.123) | (0.112) | (0.062) | (0.077) | (0.052) | (0.084) | (0.076) |
| | [0.221] | [0.321] | [0.696] | [0.226] | [0.567] | [0.404] | [0.477] |
| T2 (Learning Program + Parental IVR) | 0.222** | 0.012 | 0.101* | 0.160** | 0.083 | 0.098 | 0.117 |
| | (0.108) | (0.119) | (0.057) | (0.076) | (0.056) | (0.096) | (0.080) |
| | [0.042] | [0.917] | [0.079] | [0.038] | [0.137] | [0.308] | [0.149] |
| p-value: T1 = T2 | 0.563 | 0.318 | 0.204 | 0.417 | 0.319 | 0.046 | 0.388 |
| Observations | 6146 | 6478 | 11871 | 9312 | 12625 | 12625 | 12625 |
| Control Mean | 0.000 | 0.000 | 0.000 | 0.000 | -0.000 | -0.000 | -0.000 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 144 | 153 | 202 | 201 | 202 | 202 | 202 |

*Notes:* The table reports robustness estimates of the treatment effects on the primary student outcomes using the full endline student sample. The first specification pools the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Outcomes are standardized relative to the control group. All regressions include the same basic controls, as well as strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses. Conventional p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A22: Treatment Effects on Attrition

| | (1) Attrition |
|---|---|
| Any Treatment | -0.032* |
| | (0.018) |
| | [0.079] |
| T1 (Learning Program) | -0.055*** |
| | (0.020) |
| | [0.007] |
| T2 (Learning Program + Parental IVR) | -0.010 |
| | (0.023) |
| | [0.677] |
| *p*-value: T1 = T2 | 0.057 |
| Control Mean | 0.330 |
| Basic Controls | Yes |
| Observations | 13,383 |
| Clusters | 203 |

*Notes:* The dependent variable is an indicator for student attrition. The pooled specification combines the two treatment arms, while the subsequent specification reports treatment-arm-specific effects. All regressions include basic controls and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional *p*-values are reported in square brackets. The reported *p*-value tests equality of the T1 and T2 coefficients. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A23: Robustness to Differential Attrition: Inverse Probability Weighted (IPW) Estimates

| | **Cognitive Outcomes** | | | **Non-cognitive Outcomes** | | | |
|---|---|---|---|---|---|---|---|
| | (1) English Score (Grades 9–10) | (2) English Score (Grades 7–8) | (3) Final Marks | (4) Confidence Index | (5) Aspirations Index | (6) Gender Attitudes Index | (7) SEL Index |
| Any Treatment | 0.166** | 0.037 | 0.064 | 0.130** | 0.085* | 0.074 | 0.069 |
| | (0.077) | (0.120) | (0.057) | (0.062) | (0.047) | (0.084) | (0.068) |
| | [0.033] | [0.761] | [0.270] | [0.037] | [0.072] | [0.382] | [0.315] |
| T1 (Learning Program) | 0.130 | 0.067 | 0.010 | 0.137* | 0.054 | 0.003 | 0.019 |
| | (0.094) | (0.129) | (0.069) | (0.073) | (0.052) | (0.092) | (0.075) |
| | [0.167] | [0.605] | [0.884] | [0.064] | [0.304] | [0.977] | [0.803] |
| T2 (Learning Program + Parental IVR) | 0.203** | 0.005 | 0.118* | 0.123* | 0.118** | 0.148 | 0.120 |
| | (0.090) | (0.136) | (0.062) | (0.073) | (0.059) | (0.101) | (0.084) |
| | [0.025] | [0.969] | [0.058] | [0.092] | [0.048] | [0.143] | [0.155] |
| p-value: T1 = T2 | 0.466 | 0.585 | 0.097 | 0.859 | 0.294 | 0.134 | 0.224 |
| Observations | 4621 | 4621 | 8730 | 6867 | 9243 | 9243 | 9243 |
| p-value: T1 = T2 | 0.466 | 0.585 | 0.097 | 0.859 | 0.294 | 0.134 | 0.224 |
| Control Mean | 0.000 | -0.000 | -0.000 | -0.000 | -0.000 | -0.000 | 0.000 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Attrition IPW | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 198 | 134 | 202 | 199 | 202 | 202 | 202 |

*Notes*: The table reports IPW estimates of the effects of treatment assignment on the primary student outcomes. The weights are constructed from the estimated probability that a baseline student remains in the endline analysis sample, using pre-treatment student-, KGBV-, and warden-level characteristics, treatment indicators, and strata fixed effects. The first specification pools the two treatment arms, and the subsequent specifications report treatment-arm-specific effects. Outcomes are standardized relative to the control group. All regressions include the same basic controls as the ITT estimates and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses. Conventional p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A24: Effects on Parents' Satisfaction with the Hostel

| | **Hostel Satisfaction Outcomes** | | | | | |
|---|---|---|---|---|---|---|
| | (1) Overall Satisfaction Index | (2) Quality of Education | (3) Living Facilities | (4) Food Quality and Quantity | (5) Safety at Hostel | (6) Emotional Well-being |
| Any treatment | 0.357*** | 0.135*** | 0.143*** | 0.143*** | 0.134*** | 0.137*** |
| | (0.092) | (0.037) | (0.034) | (0.035) | (0.036) | (0.036) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| T1 (Learning Program) | 0.445*** | 0.154*** | 0.168*** | 0.168*** | 0.162*** | 0.173*** |
| | (0.105) | (0.045) | (0.042) | (0.045) | (0.044) | (0.044) |
| | [0.000] | [0.001] | [0.000] | [0.000] | [0.000] | [0.000] |
| T2 (Learning Program + Parental IVR) | 0.270*** | 0.116*** | 0.118*** | 0.119*** | 0.105** | 0.102** |
| | (0.104) | (0.042) | (0.040) | (0.040) | (0.042) | (0.039) |
| | [0.010] | [0.006] | [0.004] | [0.004] | [0.013] | [0.010] |
| p-value: T1 = T2 | 0.080 | 0.404 | 0.293 | 0.300 | 0.234 | 0.113 |
| Observations | 4216 | 4194 | 4191 | 4176 | 4195 | 4190 |
| Control Mean | 0.000 | 0.186 | 0.162 | 0.180 | 0.196 | 0.172 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 203 | 203 | 203 | 203 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on parents' overall satisfaction with the hostel and on five specific dimensions of hostel satisfaction. The first specification pools the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Treatment 1 consists of the student-level learning program only, while Treatment 2 combines the same student-level program with parental IVR messages. The overall satisfaction index in Column (1) is standardized relative to the control-group mean and standard deviation. Outcomes in Columns (2)–(6) are binary indicators, so their coefficients are interpreted as changes in probability. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

## Table A25: Effects on Parent Engagement

| | (1) Communicated with Warden (Last 10 Days) | (2) Communicated with Daughter (At Least Once/Week) | (3) Daughter Comfortable Sharing Hostel Experiences | (4) Communicated with Teacher (Last 10 Days) |
|---|---|---|---|---|
| Any treatment | 0.091* | 0.064* | 0.096** | 0.006 |
| | (0.047) | (0.039) | (0.048) | (0.039) |
| | [0.053] | [0.097] | [0.044] | [0.873] |
| T1 (Learning Program) | 0.118** | 0.095** | 0.093* | -0.001 |
| | (0.052) | (0.044) | (0.054) | (0.046) |
| | [0.024] | [0.033] | [0.085] | [0.977] |
| T2 (Learning Program + Parental IVR) | 0.063 | 0.034 | 0.100* | 0.014 |
| | (0.055) | (0.047) | (0.056) | (0.043) |
| | [0.254] | [0.475] | [0.074] | [0.750] |
| p-value: T1 = T2 | 0.302 | 0.211 | 0.892 | 0.720 |
| Observations | 4216 | 4216 | 4216 | 4216 |
| Control Mean | 0.492 | 0.626 | 0.630 | 0.278 |
| Basic Controls | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 203 | 203 |

*Notes:* The table reports OLS estimates of treatment effects on parental engagement with their daughters and the hostel. All outcomes are binary indicators. Treatment 1 consists of the student-level learning program, while Treatment 2 combines the same program with parental IVR messages. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A26: Effects on Topics Discussed with the Warden

| | (1) Studies | (2) Future Education | (3) Health | (4) Career | (5) Behaviour | (6) *Kishori Vani* |
|---|---|---|---|---|---|---|
| Any treatment | 0.013 | -0.064 | 0.015 | -0.053 | -0.011 | 0.048*** |
| | (0.024) | (0.050) | (0.031) | (0.042) | (0.049) | (0.018) |
| | [0.587] | [0.202] | [0.624] | [0.215] | [0.824] | [0.007] |
| T1 (Learning Program) | 0.047** | -0.007 | 0.020 | 0.001 | 0.001 | 0.040* |
| | (0.023) | (0.058) | (0.032) | (0.052) | (0.054) | (0.024) |
| | [0.047] | [0.907] | [0.531] | [0.990] | [0.989] | [0.094] |
| T2 (Learning Program + Parental IVR) | -0.020 | -0.121** | 0.010 | -0.105** | -0.022 | 0.055*** |
| | (0.031) | (0.059) | (0.038) | (0.046) | (0.060) | (0.019) |
| | [0.524] | [0.042] | [0.793] | [0.022] | [0.708] | [0.003] |
| p-value: T1 = T2 | 0.015 | 0.070 | 0.763 | 0.034 | 0.687 | 0.524 |
| Observations | 4216 | 4216 | 4216 | 4216 | 4216 | 4216 |
| Control Mean | 0.885 | 0.523 | 0.836 | 0.281 | 0.463 | 0.039 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 203 | 203 | 203 | 203 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on the topics discussed by parents with hostel wardens. Each outcome is a binary indicator equal to one if the parent reported discussing the specified topic with the warden and zero otherwise. The first specification pools the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Treatment 1 consists of the student-level learning program only, while Treatment 2 combines the same student-level program with parental IVR messages. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A27: Effects on Parents' Knowledge of Their Daughters and the Hostel

| | **Hostel Knowledge** | | **Knowledge of Daughter** | | | | |
|---|---|---|---|---|---|---|---|
| | (1) Knows Hostel Block | (2) Knows Warden Name | (3) Knows Favourite Subject | (4) Knows Favourite Activity | (5) Knows Student Grade | (6) Knows Educational Aspiration | (7) Knows Career Aspiration |
| Any treatment | 0.110* | 0.081* | 0.020 | 0.023 | -0.011 | 0.023 | 0.003 |
| | (0.057) | (0.049) | (0.026) | (0.025) | (0.007) | (0.025) | (0.018) |
| | [0.056] | [0.099] | [0.438] | [0.360] | [0.133] | [0.352] | [0.854] |
| T1 (Learning Program) | 0.111 | 0.055 | 0.028 | 0.029 | -0.014* | 0.044 | 0.005 |
| | (0.069) | (0.056) | (0.028) | (0.027) | (0.008) | (0.027) | (0.021) |
| | [0.109] | [0.325] | [0.319] | [0.300] | [0.073] | [0.105] | [0.805] |
| T2 (Learning Program + Parental IVR) | 0.109 | 0.107* | 0.013 | 0.017 | -0.007 | 0.002 | 0.001 |
| | (0.068) | (0.059) | (0.032) | (0.030) | (0.008) | (0.032) | (0.021) |
| | [0.109] | [0.068] | [0.692] | [0.571] | [0.345] | [0.941] | [0.943] |
| p-value: T1 = T2 | 0.984 | 0.383 | 0.611 | 0.695 | 0.267 | 0.207 | 0.858 |
| Observations | 4216 | 4216 | 4216 | 4216 | 4216 | 2876 | 2830 |
| Control Mean | 0.268 | 0.334 | 0.687 | 0.615 | 0.028 | 0.361 | 0.204 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | | |
| Clusters | 203 | 203 | 203 | 203 | 203 | 200 | 200 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on parents' knowledge of their daughters and the hostel environment. Each outcome is a binary indicator equal to one if the parent's response correctly matches the student's response or verified hostel information and zero otherwise. The first specification pools the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Treatment 1 consists of the student-level learning program only, whereas Treatment 2 combines the same student-level program with parental IVR messages. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A28: Effects on Parents' Perceptions of Gender Roles

| | (1) Cooking | (2) Leading a Team | (3) Cleaning | (4) Repairing | (5) Farming | (6) Managing Household Budget |
|---|---|---|---|---|---|---|
| Any treatment | 0.010 | 0.017 | 0.010 | -0.004 | 0.085** | 0.103** |
| | (0.040) | (0.043) | (0.040) | (0.044) | (0.040) | (0.042) |
| | [0.792] | [0.696] | [0.796] | [0.926] | [0.035] | [0.015] |
| T1 (Learning Program) | 0.002 | 0.071 | 0.030 | 0.024 | 0.112*** | 0.138*** |
| | (0.048) | (0.045) | (0.044) | (0.047) | (0.042) | (0.048) |
| | [0.967] | [0.114] | [0.489] | [0.617] | [0.009] | [0.005] |
| T2 (Learning Program + Parental IVR) | 0.019 | -0.037 | -0.010 | -0.032 | 0.058 | 0.067 |
| | (0.046) | (0.053) | (0.047) | (0.053) | (0.046) | (0.050) |
| | [0.682] | [0.492] | [0.841] | [0.551] | [0.212] | [0.180] |
| p-value: T1 = T2 | 0.739 | 0.033 | 0.383 | 0.242 | 0.165 | 0.176 |
| Observations | 4216 | 4216 | 4216 | 4216 | 4216 | 4216 |
| Control Mean | 0.766 | 0.726 | 0.723 | 0.679 | 0.761 | 0.561 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 203 | 203 | 203 | 203 |

*Notes:* The table reports OLS estimates of treatment effects on parents' perceptions of girls' ability to perform different household, agricultural, financial, and leadership roles. All outcomes are binary indicators. T1 denotes the student learning program and T2 denotes the student learning program plus parental IVR. All regressions include parent controls, KGBV controls, and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses and p-values in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A29: Effects on Parents' Support for Women's Employment

| | (1) Support City Work | (2) Support Work (100% Women) | (3) Support Work (50% Women) | (4) Support Work (33% Women) | (5) Support Work (25% Women) | (6) Support Work (10% Women) | (7) Support Work (No Women) |
|---|---|---|---|---|---|---|---|
| Any treatment | 0.026 | 0.014 | 0.038* | 0.090** | -0.010 | 0.039 | 0.018 |
| | (0.019) | (0.032) | (0.021) | (0.039) | (0.037) | (0.054) | (0.062) |
| | [0.168] | [0.664] | [0.079] | [0.023] | [0.795] | [0.473] | [0.775] |
| T1 (Learning Program) | 0.046** | 0.054* | 0.050** | 0.121*** | 0.017 | 0.015 | 0.017 |
| | (0.019) | (0.030) | (0.021) | (0.042) | (0.039) | (0.059) | (0.067) |
| | [0.014] | [0.071] | [0.020] | [0.005] | [0.672] | [0.805] | [0.799] |
| T2 (Learning Program + Parental IVR) | 0.005 | -0.026 | 0.024 | 0.054 | -0.044 | 0.075 | 0.019 |
| | (0.022) | (0.042) | (0.026) | (0.048) | (0.048) | (0.062) | (0.076) |
| | [0.803] | [0.531] | [0.364] | [0.265] | [0.364] | [0.226] | [0.805] |
| p-value: T1 = T2 | 0.014 | 0.029 | 0.217 | 0.134 | 0.201 | 0.256 | 0.981 |
| Observations | 4216 | 4216 | 3787 | 3535 | 2711 | 2117 | 1442 |
| Control Mean | 0.920 | 0.892 | 0.911 | 0.704 | 0.781 | 0.647 | 0.513 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 200 | 196 | 189 | 175 | 153 |

*Notes:* The table reports OLS estimates of the effects of treatment assignment on parents' support for girls' and women's participation in the labour market under different workplace gender compositions. Column (1) measures support for daughters working in a city, while Columns (2)–(7) measure support for daughters working in environments where women constitute different shares of the workforce. Each outcome is a binary indicator equal to one if the parent expressed support for the specified employment scenario and zero otherwise. The first specification pools the two treatment arms, while the subsequent specifications report treatment-arm-specific effects. Treatment 1 consists of the student-level learning program only, whereas Treatment 2 combines the same student-level program with parental IVR messages. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and conventional p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A30: Treatment Effects on Parents' Expectations from their Daughters

| | (1) Expected Completion Beyond Grade 12 | (2) Expected Age at Completing Education | (3) Expected Age at Marriage | (4) Expected Age at First Child | (5) Education Creates Marriage Constraints |
|---|---|---|---|---|---|
| Any treatment | 0.035 | -0.021 | 0.114 | 0.001 | 0.050 |
| | (0.026) | (0.237) | (0.238) | (0.262) | (0.032) |
| | [0.186] | [0.930] | [0.631] | [0.998] | [0.113] |
| T1 (Learning Program) | 0.077*** | 0.010 | 0.112 | 0.053 | 0.048 |
| | (0.029) | (0.255) | (0.248) | (0.267) | (0.033) |
| | [0.008] | [0.970] | [0.653] | [0.844] | [0.147] |
| T2 (Learning Program + Parent IVR) | -0.007 | -0.051 | 0.117 | -0.051 | 0.053 |
| | (0.031) | (0.287) | (0.288) | (0.320) | (0.042) |
| | [0.826] | [0.858] | [0.686] | [0.875] | [0.209] |
| p-value: T1 = T2 | 0.005 | 0.819 | 0.985 | 0.704 | 0.890 |
| Observations | 4216 | 4208 | 4208 | 4161 | 4216 |
| Control Mean | 0.745 | 22.133 | 23.294 | 24.950 | 0.117 |
| Basic Controls | Yes | Yes | Yes | Yes | Yes |
| Clusters | 203 | 203 | 203 | 203 | 203 |

*Notes:* The table reports intention-to-treat estimates of the effects of the intervention on parents' educational expectations and aspirations for their daughters. Column (1) reports the probability that parents expect their daughters to complete education beyond Grade 12. Columns (2)–(4) report parents' expected ages at completion of education, marriage, and first childbirth, respectively. Column (5) measures whether parents perceive girls' education as creating difficulties in finding a suitable groom. The first specification pools Treatment 1 and Treatment 2, while the subsequent specifications report treatment-arm-specific effects. Treatment 1 consists of the student-level learning programme, whereas Treatment 2 combines the student-level learning programme with the parent IVR intervention. All regressions include parent controls, KGBV controls, and randomization-stratum fixed effects. Robust standard errors clustered at the KGBV level are reported in parentheses, and p-values are reported in square brackets. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table A31: Spillover Effects on Newly Admitted Grade 6 Students: Cognitive Outcomes and Educational Aspirations

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | *Cognitive Outcomes* | | | *Educational Aspirations* | |
| | Final Marks | English Score | Raven Score | Study Beyond Grade 12 | University Degree Belief |
| Any Treatment | 0.077 | 0.066 | 0.077 | 0.054 | 0.054* |
| | (0.090) | (0.132) | (0.103) | (0.035) | (0.030) |
| | [0.393] | [0.618] | [0.455] | [0.130] | [0.076] |
| T1 (Lesson) | 0.053 | 0.177 | -0.116 | 0.031 | 0.066* |
| | (0.110) | (0.150) | (0.112) | (0.042) | (0.038) |
| | [0.628] | [0.240] | [0.300] | [0.461] | [0.081] |
| T2 (Lesson + IVR) | 0.098 | -0.039 | 0.258** | 0.076* | 0.043 |
| | (0.096) | (0.144) | (0.122) | (0.043) | (0.034) |
| | [0.309] | [0.790] | [0.036] | [0.076] | [0.217] |
| $p$-value: T1 = T2 | 0.656 | 0.108 | 0.003 | 0.329 | 0.556 |
| Control Mean | 0.000 | -0.000 | 0.000 | 0.699 | 0.712 |
| Observations | 2,497 | 2,732 | 2,732 | 2,571 | 2,616 |
| Clusters | 133 | 134 | 134 | 134 | 134 |
| Student Controls | Yes | Yes | Yes | Yes | Yes |
| KGBV Controls | Yes | Yes | Yes | Yes | Yes |
| Warden Controls | Yes | Yes | Yes | Yes | Yes |
| Strata FE | Yes | Yes | Yes | Yes | Yes |

*Notes:* The table reports OLS estimates of spillover effects on newly admitted Grade 6 students who did not directly receive the intervention. "Any Treatment" pools the two treatment arms. T1 denotes the student learning intervention, while T2 combines the student learning intervention with the parental IVR intervention. Cognitive outcomes are standardized relative to the Grade 6 control-group mean and standard deviation. All regressions include student-level controls, KGBV-level controls, warden-level controls, and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and $p$-values are reported in square brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A32: Spillover Effects on Newly Admitted Grade 6 Students: Socio-Emotional Learning Outcomes

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | *Socio-Emotional Learning Outcomes* | | | | |
| | SEL Index | Self-Awareness | Social Awareness | Self-Management | Anxiety |
| Any Treatment | 0.079 | -0.016 | 0.137 | 0.040 | 0.017 |
| | (0.116) | (0.113) | (0.113) | (0.098) | (0.120) |
| | [0.496] | [0.886] | [0.228] | [0.687] | [0.886] |
| T1 (Lesson) | 0.151 | 0.002 | 0.222* | 0.101 | 0.048 |
| | (0.135) | (0.122) | (0.129) | (0.122) | (0.143) |
| | [0.267] | [0.986] | [0.087] | [0.411] | [0.739] |
| T2 (Lesson + IVR) | 0.014 | -0.033 | 0.059 | -0.016 | -0.011 |
| | (0.128) | (0.127) | (0.127) | (0.110) | (0.137) |
| | [0.913] | [0.796] | [0.641] | [0.883] | [0.938] |
| $p$-value: T1 = T2 | 0.278 | 0.736 | 0.180 | 0.348 | 0.686 |
| Control Mean | -0.000 | 0.000 | -0.000 | -0.000 | -0.000 |
| Observations | 2,732 | 2,732 | 2,732 | 2,732 | 2,732 |
| Clusters | 134 | 134 | 134 | 134 | 134 |
| Student Controls | Yes | Yes | Yes | Yes | Yes |
| KGBV Controls | Yes | Yes | Yes | Yes | Yes |
| Warden Controls | Yes | Yes | Yes | Yes | Yes |
| Strata FE | Yes | Yes | Yes | Yes | Yes |

*Notes:* The table reports OLS estimates of spillover effects on socio-emotional learning outcomes among newly admitted Grade 6 students who did not directly receive the intervention. "Any Treatment" pools the two treatment arms. T1 denotes the student learning intervention, while T2 combines the student learning intervention with the parental IVR intervention. Outcomes are standardized relative to the Grade 6 control-group mean and standard deviation. All regressions include student-level controls, KGBV-level controls, warden-level controls, and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and $p$-values are reported in square brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A33: Spillover Effects on Newly Admitted Grade 6 Students: Gender Attitudes

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *Gender Attitudes* | | | |
| | Gender Attitude Index | Education Attitudes | Employment Attitudes | Equal Rights Attitudes |
| Any Treatment | 0.005 | 0.325*** | 0.048 | 0.105 |
| | (0.129) | (0.096) | (0.093) | (0.099) |
| | [0.970] | [0.001] | [0.608] | [0.294] |
| T1 (Lesson) | -0.159 | 0.375*** | 0.012 | 0.206* |
| | (0.130) | (0.106) | (0.101) | (0.115) |
| | [0.224] | [0.001] | [0.907] | [0.075] |
| T2 (Lesson + IVR) | 0.154 | 0.279** | 0.080 | 0.012 |
| | (0.151) | (0.113) | (0.110) | (0.113) |
| | [0.311] | [0.015] | [0.465] | [0.913] |
| $p$-value: T1 = T2 | 0.015 | 0.364 | 0.505 | 0.096 |
| Control Mean | -0.000 | 0.000 | 0.000 | -0.000 |
| Observations | 2,732 | 2,732 | 2,732 | 2,732 |
| Clusters | 134 | 134 | 134 | 134 |
| Student Controls | Yes | Yes | Yes | Yes |
| KGBV Controls | Yes | Yes | Yes | Yes |
| Warden Controls | Yes | Yes | Yes | Yes |
| Strata FE | Yes | Yes | Yes | Yes |

*Notes:* The table reports OLS estimates of spillover effects on gender attitudes among newly admitted Grade 6 students who did not directly receive the intervention. Column (1) reports the overall Gender Attitude Index, while Columns (2)–(4) report its components relating to education, employment, and equal rights. "Any Treatment" pools the two treatment arms. T1 denotes the student learning intervention, while T2 combines the student learning intervention with the parental IVR intervention. Outcomes are standardized relative to the Grade 6 control-group mean and standard deviation. All regressions include student-level controls, KGBV-level controls, warden-level controls, and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and $p$-values are reported in square brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A34: Spillover Effects on Siblings of KGBV Students

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | *Cognitive Outcomes* | | *Awareness and Attitudes* | | *Educational Aspirations* | |
| | Raven's Puzzle | ASER Score | Gender Attitude | General Awareness | University Aspiration | University Ability |
| Any Treatment | 0.128 | 0.055 | 0.126 | 0.226 | 0.133*** | 0.053 |
| | (0.124) | (0.087) | (0.128) | (0.137) | (0.038) | (0.048) |
| | [0.304] | [0.529] | [0.327] | [0.101] | [0.001] | [0.268] |
| T1 (Lesson) | 0.153 | 0.068 | 0.103 | 0.285** | 0.155*** | 0.095* |
| | (0.137) | (0.098) | (0.130) | (0.136) | (0.038) | (0.057) |
| | [0.269] | [0.490] | [0.428] | [0.038] | [0.000] | [0.097] |
| T2 (Lesson + IVR) | 0.103 | 0.042 | 0.149 | 0.166 | 0.111** | 0.009 |
| | (0.160) | (0.117) | (0.163) | (0.185) | (0.049) | (0.055) |
| | [0.523] | [0.722] | [0.362] | [0.369] | [0.024] | [0.866] |
| $p$-value: T1 = T2 | 0.763 | 0.835 | 0.755 | 0.490 | 0.298 | 0.161 |
| Control Mean | -0.011 | -0.007 | -0.013 | -0.011 | 0.682 | 0.564 |
| Observations | 1,287 | 1,280 | 1,287 | 1,287 | 1,287 | 1,287 |
| Clusters | 158 | 158 | 158 | 158 | 158 | 158 |
| Sibling Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Parent/Household Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Strata FE | Yes | Yes | Yes | Yes | Yes | Yes |

*Notes:* The table reports OLS estimates of spillover effects on siblings of students enrolled in KGBV hostels. "Any Treatment" pools the two treatment arms. T1 denotes the student learning intervention, while T2 combines the student learning intervention with the parental IVR intervention. All regressions include sibling-level controls, parent/household controls, and strata fixed effects. Standard errors clustered at the KGBV level are reported in parentheses, and $p$-values are reported in square brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

# Supplementary Appendix

Table S1: Content of the English Lessons and Related Skills

| No. | Main Content | Related Skills |
|---|---|---|
| 1. | Phonic sounds and alphabet | Letter-sound correspondence, recognition and writing of English letters, basic decoding of words |
| 2. | CVC words and word families | Blending and segmenting sounds, reading and spelling of simple CVC words, recognizing common spelling patterns |
| 3. | Sight words and basic vocabulary | Recognition of high-frequency words, development of everyday vocabulary, support for reading fluency |
| 4. | English instructions and classroom language | Listening to and understanding simple instructions, following directions, using functional classroom English |
| 5. | Greetings and self-introduction | Introducing oneself and others, asking and answering simple Wh-questions, conducting short conversations |
| 6. | Parts of the body, actions and daily routines | Thematic vocabulary for body parts and actions, use of common verbs, describing personal routines and preferences |
| 7. | Grammar fundamentals | Identification and use of nouns, verbs and pronouns; constructing simple sentences with correct word order |
| 8. | Reading comprehension | Understanding short texts and passages, locating key information, answering comprehension questions, making simple inferences |
| 9. | Writing practice | Writing letters and words, forming short sentences, applying basic punctuation and spelling |
| 10. | Songs, games and role-play | Listening and pronunciation practice, rhythm and stress, building confidence in speaking, cooperative interaction in pairs and groups |

Table S2: Content of Life-skills

| No. | Main Content | Related Skills |
|---|---|---|
| 1. | Socio-Emotional Learning (SEL) | Self-awareness, emotional regulation, interpersonal skills |
| 2. | Health, Hygiene, and Nutrition | Menstrual health, personal hygiene, nutritional awareness |
| 3. | Gender Sensitivity and Stereotype | Understanding gender roles, questioning stereotypes, confidence-building |
| 4. | Personal Safety and Boundaries | Recognizing safe and unsafe touch, self-protection, seeking support |
| 5. | Decision-Making and Goal Setting | Critical thinking, problem-solving, setting and achieving goals |
| 6. | Social Awareness and Empathy | Empathy, community engagement, leadership skills |

Table S3: Parent IVRS Content

| Weeks | Module | Content |
|---|---|---|
| 1–3 | Programme Introduction | Introduction to Kishori Vani IVRS calls, programme purpose, and initial parent feedback |
| 4–8 | Parental Engagement and Girls' Education | Supporting girls' aspirations; parental involvement in education; educational outcomes; school retention and continuation to higher education |
| 9–10 | IVRS Awareness and Parent Feedback | Awareness of the IVRS programme; parent perceptions of the calls; feedback on content and communication preferences |
| 11–15 | Education, Family Expectations, and Future Opportunities | Value of education relative to early marriage and work; managing family expectations; community support for girls' education; perceptions of labour-market opportunities |
| 16 | Parent Contact Verification | Verification of parent contact information for continued IVRS delivery |
| 17–21 | Gender Norms and Stereotypes | Local female role models; identifying and challenging gender stereotypes; promoting equality at home; sharing household responsibilities; supporting girls' autonomy |
| 22–26 | Girls' Rights, Work, and Leadership | Girls' participation in decisions; education, career, and marriage choices; women's right to work; economic independence; women's leadership |
| 27–30 | Child Marriage and Child Labour | Consequences of child marriage; legal age of marriage; identifying child labour; consequences for education, skills, and future opportunities; recap and quiz |
| 31–35 | Scholarships and Financial Support for Girls | Scholarships for girls' education and higher education; credible information sources; government savings and scholarship schemes; parent awareness and recap |
| 36–39 | Skills, Vocational Training, and Internships | Importance of vocational skills; skill-development opportunities; government training programmes; internships and pathways into employment |
| 40–44 | Careers, Employment, and Support Services | Formal and informal employment; education and career pathways; challenging occupational stereotypes; women's helpline services; recap and parent engagement check |
| 45–48 | Mental Health and Emotional Well-Being | Understanding mental health; gender and emotional needs; parent–daughter communication; managing emotions and stress |
| 49–52 | Positive Thinking, Confidence, and Self-Reliance | Positive self-talk; responding constructively to negative thoughts; building confidence; encouraging decision-making, initiative, and self-reliance |

*Notes:* The table groups the 52 weekly parent IVRS messages into broad content modules for ease of presentation. Module labels consolidate closely related weekly messages. In addition to substantive content, selected weeks included programme introduction, parent feedback, contact verification, and short quizzes or engagement checks.

# S1 Construction of Cognitive Outcomes

## S1.1 English Score

Students completed seven multiple-choice questions covering vocabulary, spelling, grammar, translation, sentence construction, and reading comprehension. We developed two test forms: one for the younger grade group and a more difficult version for the older grade group. Each correct response was awarded one point, and the English test score was calculated as the total number of correct answers. We then standardized the score using the control-group mean and standard deviation within the relevant test form.

Table S4: English Assessment Items (Grades 9-10)

| No. | Assessment item and response options |
|---|---|
| Q125 | How many consonants are in the word "SUNFLOWER"?<br>A. 4<br>B. 6<br>C. 5<br>D. 3 |
| Q126 | What does the word "BRAVE" mean?<br>A. Clever<br>B. Fearless<br>C. Angry<br>D. Quiet |
| Q127 | The options below include the names of seasons. Identify the correctly spelled one.<br>A. Wintar<br>B. Monson<br>C. Summer<br>D. Autum |
| Q128 | Translate the following into English:<br>ನಾನು ನನ್ನ ಸ್ನೇಹಿತರೊಂದಿಗೆ ಆಟ ಆಡುವುದನ್ನು ಇಷ್ಟಪಡುತ್ತೇನೆ.<br>A. I like my friends.<br>B. I like to play with my friends.<br>C. I like my classmates.<br>D. I like to talk with my friends. |

Continued on next page

*Table S4 continued*

| No. | Assessment item and response options |
|---|---|
| Q129 | Fill in the blank with the correct word:<br>My father _____ the newspaper every morning.<br>A. read<br>B. reads<br>C. reading<br>D. readed |
| Q130 | Identify the correct sentence.<br>A. The garden plays she in.<br>B. She in the garden plays.<br>C. In the garden she play.<br>D. She plays in the garden. |
| Q131 | Read the passage and answer the question below.<br>*Riya is my cousin. She loves chocolate. We live in the same village. It is Sunday today. She is wearing a blue dress. She looks very happy.*<br>*How is Riya looking today?*<br>A. Dress<br>B. Happy<br>C. Very<br>D. Blue |

Table S5: English Assessment Items (Grades 7-8)

| No. | Assessment item and response options |
|---|---|
| Q125 | How many vowels are in the word “SUNFLOWER”?<br>A. 4<br>B. 6<br>C. 5<br>D. 3 |
| Q126 | What is the opposite of the word “failure”?<br>A. Defeat<br>B. Success<br>C. Weakness<br>D. Collapse |

Continued on next page

*Table S5 continued*

| No. | Assessment item and response options |
|---|---|
| Q127 | The options below include the names of months. Identify the correctly spelled one.<br>A. Septamber<br>B. Septembar<br>C. September<br>D. Saptember |
| Q128 | Identify the correct sentence.<br>A. An red apple is.<br>B. Apple an is red.<br>C. An apple is red.<br>D. Red is an apple. |
| Q129 | Fill in the blank with the correct word:<br>She _____ to the market yesterday.<br>A. Go<br>B. Went<br>C. Going<br>D. Goes |
| Q130 | Which word is similar in meaning to the word "joyful"?<br>A. Healthy<br>B. Proud<br>C. Happy<br>D. Strong |
| Q131 | Read the passage and answer the question below.<br>*Riya is my cousin. She loves chocolates. We live in the same town. It is Sunday today. She is wearing a blue dress. She looks very happy.*<br>*What does Riya love?*<br>A. Cousin<br>B. Sunday<br>C. Chocolates<br>D. Dress |

# S2 Construction of Non-Cognitive Outcomes

## S2.1 Confidence Index

The Confidence Index measures students' confidence using three incentivized behavioral tasks. It captures three complementary dimensions of confidence: *leadership confidence*, *public confidence*, and *private confidence*. Unlike conventional self-reported measures, these tasks involve consequential choices and are designed to elicit students' behavioral expressions

of confidence. The composite index is constructed from the following three components, with higher values indicating greater confidence.

1. *Leadership Confidence*: Students were asked whether they would like to become the class monitor. Before making this decision, they were informed that one volunteer would be selected at random and required to deliver a one-minute speech explaining why she would be a suitable candidate. Leadership confidence is coded as 1 if the student volunteered to become class monitor and 0 otherwise. The indicator is rescaled to range from 0 to 100. If selected, a student received 3 gift tokens for delivering the speech adequately, 1 token for making only a limited attempt, and no token if she refused or did not speak.

2. *Public Confidence*: Students chose whether to solve an additional Raven puzzle publicly on the board or privately on their own paper. Before making this choice, they were informed that one student among those who selected the public option would be chosen at random to solve the puzzle in front of the class. Public confidence is coded as 1 if the student chose to solve the puzzle publicly and 0 if she chose to solve it privately. This indicator is also rescaled to range from 0 to 100. A selected student received 3 gift tokens for solving the puzzle correctly and 1 token for attempting the task but providing an incorrect answer.

3. *Private Confidence*: Students were asked to predict how many questions they had answered correctly in the Raven reasoning and English tests before receiving any information about their actual performance. They were informed that they would receive gift tokens based on the accuracy of their predictions. The two tests contained 5 Raven questions and 7 English questions, giving a total of 12 items. The private-confidence measure is calculated as the total number of questions the student expected to have answered correctly, divided by 12 and multiplied by 100.

## S2.2 Aspirations Index

We adapt from Beaman et al. (2012) to construct the Aspirations Index, which captures students' aspirations regarding future education, occupation, and marriage. Higher values indicate stronger educational and occupational aspirations and a preference for delaying marriage. The index is constructed from the following three binary components:

1. *Study Beyond Grade 12*: This indicator is constructed from Q27 and equals 1 if the student aspires to complete either a university degree (Bachelor's or Master's) or a diploma, such as a polytechnic or nursing qualification, and 0 otherwise.

2. *Professional Occupation*: This indicator is constructed from Q33 and equals 1 if the student aspires to become a doctor, engineer, government officer, police or defense officer, or politician, and 0 otherwise.

3. *Marriage After Age 25*: This indicator is constructed from Q34 and equals 1 if the student aspires to marry at age 25 or later, or reports being too young to think about marriage, and 0 otherwise.

## S2.3 Gender Attitudes and Social Norms

We adapt from Dhar et al. (2022) to create the Gender Attitudes Index, related domain-specific sub-indices, and the Social Norms Index.

Gender Attitudes Index is constructed using students' responses to 9 statements on gender roles and equality (Q40–48) in the endline survey. These statements cover attitudes toward women's roles within the household, educational opportunities for girls, equal rights, political leadership, and gender-based violence. Survey items were recoded where necessary so that higher values consistently indicate more progressive gender attitudes. The index was then constructed as the first principal component obtained from a principal component analysis (PCA) of the recoded items.

The Social Norms Index captures students' perceptions of prevailing community norms regarding women's education, employment, and gender roles. It is based on responses to items Q52–Q56. Survey items were recoded where necessary so that higher values indicate more progressive social norms. The index was constructed as the first principal component obtained from a PCA of the recoded items.

We also construct three thematic sub-indices using the same procedure:

1. *Education Attitudes Index*: Items Q43, Q44, Q46, and Q54, capturing attitudes toward girls' educational opportunities and educational attainment.

2. *Employment Attitudes Index*: Items Q40 and Q52, capturing attitudes toward women's economic roles and labor-force participation.

3. *Equal Rights Attitudes Index*: Items Q41, Q42, Q45, and Q47, capturing beliefs about women's decision-making authority, tolerance of violence, and equal legal and social rights.

Table S6: Survey Items Used to Construct the Gender Attitudes and Social Norms Index

| No. | Survey item | Response options |
|---|---|---|
| Q40 | A woman's most important role is being a good homemaker (feeding kids and cooking for her family). | A. Strongly agree<br>B. Agree<br>C. Neither agree nor disagree<br>D. Disagree<br>E. Strongly disagree |
| Q41 | A man should have the final word about decisions in his home. | |
| Q42 | A woman should tolerate violence to keep her family together. | |
| Q43 | Wives should be less educated than their husbands. | |
| Q44 | Boys should get more opportunities/resources for education. | |
| Q45 | Men and women should get equal opportunities in all spheres of life. | |
| Q46 | Girls should be allowed to study as far as they want. | |
| Q47 | Daughters should have a similar right to inherited property as sons. | |
| Q48 | It would be a good idea to elect a woman as the village Sarpanch. | |
| Q52 | Women should be allowed to work outside home. | A. Strongly agree<br>B. Agree<br>C. Neither agree nor disagree<br>D. Disagree<br>E. Strongly disagree |
| Q53 | People in your village/community think that women should be allowed to work outside home. | |
| Q54 | Girls should be allowed to study in college even if it is far away. | |
| Q55 | People in your village/community think that girls should be allowed to study in college even if it is far away. | |
| Q56 | The community will oppose you if you disagree with them. | |

## S2.4 Social and Emotional Learning (SEL) Index

The Social and Emotional Learning (SEL) Index is a 12-item (Q77-Q88) survey adapted from Anthony et al. (2020) and Ganicho et al. (2025) that measures students' socio-emotional competencies across three broad domains: self-awareness, self-management, and social awareness.

Table S7: Survey Items Used to Construct the SEL Index

| No | Survey item | Response options |
|---|---|---|
| Q77 | I ask for help when I need it. | A. Never<br>B. Rarely<br>C. Often<br>D. Always |
| Q78 | I do my part in a group. | |
| Q79 | I let people know when there is a problem. | |
| Q80 | I pay attention when others present their ideas. | |
| Q81 | I stay calm when dealing with problems. | |
| Q82 | I stay calm when I disagree with others. | |
| Q83 | I pay attention when the teacher talks to the class. | |
| Q84 | I try to find a good way to end a disagreement. | |
| Q85 | I help my friends when they are having a problem. | |
| Q86 | I stand up for others when they are not treated well. | |
| Q87 | I try to make others feel better. | |
| Q88 | I try to think about how others feel. | |

## S2.5 Perceived Barriers Index

The Perceived Barriers Index is constructed from four survey questions (Q29–Q32), which capture students' perceptions of obstacles to achieving their educational aspirations. It captures the extent to which students perceive gender, caste, household financial constraints, and family attitudes as barriers to pursuing further education. Survey items are recoded so that higher values indicate fewer perceived barriers to education.

Table S8: Survey Items Used to Construct the Perceived Barriers Index

| No | Survey item | Response options |
|---|---|---|
| Q29 | Do you think that because you are a girl, it could be an obstacle to the achievement of your educational aim? | A. Very much<br>B. Somewhat<br>C. Very little<br>D. Not at all |
| Q30 | Do you think caste prejudice could be an obstacle to the achievement of your educational aim? | |
| Q31 | Do you think a lack of money in your family could be an obstacle to the achievement of your educational aim? | |
| Q32 | Do you think the ideas of your family could be an obstacle to the achievement of your educational aim? | |

## S2.6 Education Engagement:

Students' education engagement is measured using five survey questions (Q2 and Q7–Q10). Students reported the number of school days they had missed during the previous week. We construct a binary indicator, *Never Missed School*, equal to 1 if the student reported no absence during the previous week and 0 otherwise. Students also reported their private study time in the morning and evening on school days, during weekends, and on days spent at home. From these responses, we construct three outcomes measured in hours: study time on school days during a typical week, total weekly study time in the hostel, and weekly study time while visiting home.

# S3 Construction of Parent-Level Outcome Variables

## S3.1 Importance of Education Index

Parents' perceived importance of education is measured using four survey questions (Q36–Q39). Parents reported how likely they believed completing a diploma or college degree would affect their daughter's life across four dimensions: future success, employment and earnings, social status, and financial independence. Responses were recorded on a five-point Likert scale ranging from 1 (*Not at all likely*) to 5 (*Very likely*). For each item, we construct

Table S9: Survey Items Used to Construct Education Engagement Outcomes

| No | Survey item | Response options |
|---|---|---|
| Q2 | In the past week, how many days were you absent from school? | A. Never<br>B. 1 day<br>C. 2 days<br>D. 3 days<br>E. 4 days<br>F. 5 days or full week |
| Q7 | On a typical school-day morning, how much time do you usually study in the hostel before going to school? | A. 0 hours<br>B. Up to 0.5 hours<br>C. 0.5–1 hour<br>D. 1–1.5 hours<br>E. 1.5–2 hours<br>F. More than 2 hours |
| Q8 | On a typical school-day evening, how much time do you usually study in the hostel after returning from school? | A. 0 hours<br>B. Up to 1 hour<br>C. 1–2 hours<br>D. 2–3 hours<br>E. 3–4 hours<br>F. More than 4 hours |
| Q9 | On a typical weekend (Saturday and Sunday), how much time do you usually study in the hostel? | A. 0 hours<br>B. Up to 2 hours<br>C. 2–4 hours<br>D. 4–6 hours<br>E. 6–8 hours<br>F. 8–10 hours<br>G. More than 10 hours |
| Q10 | On a typical day, how much time do you usually study when you are at home? | A. 0 hours<br>B. Up to 1 hour<br>C. 1–2 hours<br>D. 2–3 hours<br>E. 3–4 hours<br>F. More than 4 hours |

a binary indicator equal to 1 if the parent responded *Very likely* and 0 otherwise. These indicators are used to construct the *Importance of Education Index*.

Table S10: Survey Items Used to Construct the Importance of Education Index

| No | Survey item | Response options |
|---|---|---|
| Q36 | How much do you believe that completing a diploma or college degree will impact your daughter's future success? | 1. Not at all likely<br>2. Unlikely<br>3. Somewhat likely<br>4. Likely<br>5. Very likely |
| Q37 | How much do you believe that completing a diploma or college degree will impact your daughter's ability to secure a well-paying job in the future and earn more than your current income? | 1. Not at all likely<br>2. Unlikely<br>3. Somewhat likely<br>4. Likely<br>5. Very likely |
| Q38 | How much do you believe that completing a diploma or college degree will impact your daughter's social status in the community? | 1. Not at all likely<br>2. Unlikely<br>3. Somewhat likely<br>4. Likely<br>5. Very likely |
| Q39 | How much do you believe that completing a diploma or college degree will contribute to your daughter's financial independence in the future? | 1. Not at all likely<br>2. Unlikely<br>3. Somewhat likely<br>4. Likely<br>5. Very likely |

## S3.2 Perceived Returns to Education

Parents' perceived returns to education are measured using four outcomes from survey question Q42. Parents reported the monthly earnings, in Rupees, that they expected their daughter could earn if she completed Grade 8, Grade 10, Grade 12, and college, respectively. We construct four measures of expected monthly earnings corresponding to these levels of educational attainment. The reported earnings are converted to logarithmic values and used as outcomes in the analysis.

## S3.3 Confidence in Daughter's Educational and Professional Attainment

Parents' confidence in their daughter's educational and professional attainment is measured using two survey questions (Q30 and Q33). Parents reported how likely they believed their daughter would be able to study until the level or age she desired and how likely she would be able to achieve the career level to which she aspired. Responses were recorded on a five-point Likert scale ranging from 1 (*Not at all likely*) to 5 (*Very likely*). We construct

Table S11: Survey Items Used to Construct Perceived Returns to Education Outcomes

| No | Survey item | Response options |
|---|---|---|
| Q42.1 | How much do you think your daughter could earn in a month if she completes Grade 8? | Monthly earnings in Rupees |
| Q42.2 | How much do you think your daughter could earn in a month if she completes Grade 10? | Monthly earnings in Rupees |
| Q42.3 | How much do you think your daughter could earn in a month if she completes Grade 12? | Monthly earnings in Rupees |
| Q42.4 | How much do you think your daughter could earn in a month if she completes college? | Monthly earnings in Rupees |

two binary outcomes: *Confidence in Educational Attainment* and *Confidence in Professional Attainment*. For each outcome, the indicator is equal to 1 if the parent responded *Very likely* and 0 otherwise.

Table S12: Survey Items Used to Construct Confidence in Daughter's Attainment Outcomes

| No | Survey item | Response options |
|---|---|---|
| Q30 | On a scale of 1 to 5, how likely do you think your daughter will be able to study until the level or age she desires? | 1. Not at all likely<br>2. Unlikely<br>3. Somewhat likely<br>4. Likely<br>5. Very likely |
| Q33 | On a scale of 1 to 5, how likely do you think your daughter will be able to achieve the career level she aspires to? | 1. Not at all likely<br>2. Unlikely<br>3. Somewhat likely<br>4. Likely<br>5. Very likely |

## S3.4 Expectations for Daughter's Life Course

Parents' expectations for their daughter's life course are measured using five survey questions (Q40 and Q43–Q46). Parents reported when they expected their daughter to leave full-time education, get married, and have her first child. They also reported the level of education they believed their daughter should complete and whether they believed that her education could make it more difficult to find a groom. The first three outcomes capture parents' expectations regarding the timing of key life events. For educational expectation about daughter, we construct a binary indicator, *Education Beyond Grade 12*, equal to 1 if the parent reported that the daughter should complete a university degree and 0 otherwise. We also construct a binary indicator, *Education Makes Finding a Groom Difficult*, equal to 1 if the parent believed that the daughter's education could make it difficult to find a groom and 0 otherwise.

Table S13: Survey Items Used to Construct Expectations for Daughter's Life Course Outcomes

| No | Survey item | Response options |
|---|---|---|
| Q43 | When do you expect your daughter to leave full-time education? | Expected timing of leaving full-time education |
| Q44 | When do you expect your daughter to get married? | Expected timing of marriage |
| Q45 | When do you expect your daughter to have a child? | Expected timing of first childbirth |
| Q46 | To what level should she complete her studies? | 1. She should not study anymore after this grade<br>2. SSLC (up to Class X)<br>3. PUC (up to Class XII)<br>4. Diploma (polytechnic/nursing, etc.)<br>5. University Degree (Bachelor/Master) |
| Q40 | Do you think your girl's education could make it difficult to find a groom for your daughter's marriage? | 1. Yes<br>2. No |

## S3.5 Parental Engagement with Daughter and Warden/School

Parental engagement with the daughter and Warden/School is measured using four survey questions. Parents reported the frequency of household communication with their daughter while she was residing in the hostel, the most recent communication between someone in the household and the hostel warden and school teacher, and their daughter's comfort in sharing hostel experiences and problems. We construct a binary indicator, *Communicated with Daughter(At least Once/week)*, equal to 1 if someone in the household communicated with the daughter almost every day or at least once a week and 0 otherwise. We construct two additional binary indicators, *Communicated with Warden(Last 10 days)* and *Communicated with Teacher*, equal to 1 if someone in the household had communicated with the warden or teacher, respectively, within the previous 10 days and 0 otherwise. Finally, we construct a binary indicator, *Daughter Comfortable Sharing Hostel Experiences*, equal to 1 if the parent reported that the daughter was comfortable or very comfortable sharing her hostel experiences and problems and 0 otherwise.

## S3.6 Hostel Satisfaction

Parents' satisfaction with their daughter's hostel is measured using five survey questions (Q17–Q21). Parents rated their satisfaction with the quality of education, living facilities, food quality and quantity, their daughter's safety at the hostel, and her emotional well-being. Responses were recorded on a five-point Likert scale ranging from 1 (*Not at all satisfied*) to

Table S14: Survey Items Used to Construct Parental Engagement Outcomes

| No | Survey item | Response options |
|---|---|---|
| Q23 | How often does anyone in the household communicate with your daughter over the phone while she is in the hostel? | 1. Almost every day<br>2. At least once a week<br>3. At least once every two weeks<br>4. At least once a month<br>5. Less than once a month |
| Q14 | When was the last time someone from your household had any form of communication with the hostel/warden? | 1. 10 days or less than 10 days ago<br>2. 20–30 days ago<br>3. 1–2 months ago<br>4. 3–4 months ago<br>5. 6 months ago<br>6. More than 6 months ago |
| Q16 | When was the last time someone from the household had any form of communication with the school teacher? | 1. 10 days or less than 10 days ago<br>2. 20–30 days ago<br>3. 1–2 months ago<br>4. 3–4 months ago<br>5. 6 months ago<br>6. More than 6 months ago<br>7. Never spoken |
| Q21 | On a scale of 1 to 5, please rate how comfortable your daughter is in sharing her hostel experiences and problems. | 1. Not at all comfortable<br>2. Slightly comfortable<br>3. Moderately comfortable<br>4. Comfortable<br>5. Very comfortable |

5 (*Very satisfied*). We combine the five measures using principal component analysis (PCA) to construct an *Overall Satisfaction Index*. The resulting index is standardized using the mean and standard deviation of the control group, such that higher values indicate greater parental satisfaction with the daughter's hostel.

Table S15: Survey Items Used to Construct the Hostel Satisfaction Index

| **No** | **Survey item** | **Response options** |
|---|---|---|
| Q17 | How satisfied are you with the quality of education at your daughter's KGBV hostel? | 1. Not at all satisfied<br>2. A little satisfied<br>3. Somewhat satisfied<br>4. Satisfied<br>5. Very satisfied |
| Q18 | How satisfied are you with the standard of living facilities at your daughter's KGBV hostel? | 1. Not at all satisfied<br>2. A little satisfied<br>3. Somewhat satisfied<br>4. Satisfied<br>5. Very satisfied |
| Q19 | How satisfied are you with the food quantity and quality at your daughter's KGBV hostel? | 1. Not at all satisfied<br>2. A little satisfied<br>3. Somewhat satisfied<br>4. Satisfied<br>5. Very satisfied |
| Q20 | How satisfied are you with your daughter's safety at the KGBV hostel? | 1. Not at all satisfied<br>2. A little satisfied<br>3. Somewhat satisfied<br>4. Satisfied<br>5. Very satisfied |
| Q21 | How satisfied are you with your daughter's emotional well-being at the KGBV hostel? | 1. Not at all satisfied<br>2. A little satisfied<br>3. Somewhat satisfied<br>4. Satisfied<br>5. Very satisfied |

## S3.7 Parental Knowledge of Daughter and Hostel

Parents' knowledge of their daughter and her hostel is measured using seven outcomes: knowledge of the hostel block, warden's name, daughter's favourite subject and activity, grade, educational aspiration, and career aspiration. For the hostel block and warden's name, parent responses are compared with administrative information available to the research team. For the remaining five outcomes, parent responses are compared with the corresponding information reported by their daughter in the student survey. For each outcome, we construct a binary indicator equal to 1 if the parent's response matches the corresponding administrative or student-reported information and 0 otherwise. The seven outcomes are retained separately in the analysis.

Table S16: Survey Items Used to Construct Parental Knowledge Outcomes

| No | Survey item | Matching information |
|---|---|---|
| Q12 | What is the name of the KGBV/KKGBV block where your girl child is staying? | Administrative records |
| Q13 | What is the name of the hostel warden? | Administrative records |
| Q25 | What is your daughter's favourite subject in school? | Student-reported favourite subject |
| Q26 | Other than studying, what else does she enjoy doing? | Student-reported favourite activity |
| Q27 | How many marks (or grade) in total did your daughter get in last year's exam? | Student-reported grade |
| Q28 | What level of higher education does your daughter want to study in the future? | Student-reported educational aspiration |
| Q32 | What does your daughter want to become in the future? | Student-reported career aspiration |

## S3.8 Topics Discussed with the Warden

Topics discussed with the warden are measured using survey question Q15. Parents reported the topics that someone from the household usually discussed with the warden, with multiple responses allowed. We construct six binary outcomes indicating whether the discussion included the daughter's current studies, future education plans, physical health, career or job possibilities, behaviour in school or the hostel, and Kishori Vani. Each indicator is equal to 1 if the parent reported discussing the corresponding topic and 0 otherwise. The six outcomes are retained separately in the analysis.

Table S17: Survey Items Used to Construct Topics Discussed with the Warden Outcomes

| No | Survey item | Response options |
|---|---|---|
| Q15.1 | Current studies | 1. Yes<br>0. No |
| Q15.2 | Future education plans | 1. Yes<br>0. No |
| Q15.3 | Physical health | 1. Yes<br>0. No |
| Q15.4 | Career or job possibilities | 1. Yes<br>0. No |
| Q15.5 | Behaviour in school or hostel | 1. Yes<br>0. No |
| Q15.6 | Kishori Vani (IVRS) | 1. Yes<br>0. No |

*Notes:* Q15 allowed multiple responses. Each topic is coded as a separate binary indicator equal to 1 if the parent reported discussing the corresponding topic with the warden and 0 otherwise.

## S3.9 Parents' Support for Women's Employment

Parents' support for their daughter's future employment is measured using seven outcomes (Q35 and Q72). Parents were asked whether they would allow their daughter to work in a workplace with varying shares of female workers, ranging from an all-women workplace to a workplace with no female workers. For each workplace composition, we construct a separate binary indicator equal to 1 if the parent reported that they would allow their daughter to work in such a workplace and 0 otherwise. We additionally construct a binary indicator, *Support for Work in Another City*, equal to 1 if the parent reported that they would support their daughter working in another city, provided it was safe, and 0 if they believed that girls should remain close to home. The seven outcomes are retained separately in the analysis.

Table S18: Survey Items Used to Construct Parents' Support for Women's Employment Outcomes

| **No** | **Survey item** | **Response options** |
|---|---|---|
| Q72 | Would you support your daughter if she wants to work in another city? | 1. Yes, if it is safe<br>0. No, girls should stay close to home |
| Q35.1 | Would you allow your daughter to work in a workplace where all workers are women (100%)? | 1. Yes<br>0. No |
| Q35.2 | Would you allow your daughter to work in a workplace where half of the workers are women (50%)? | 1. Yes<br>0. No |
| Q35.3 | Would you allow your daughter to work in a workplace where one-third of the workers are women (33%)? | 1. Yes<br>0. No |
| Q35.4 | Would you allow your daughter to work in a workplace where one-fourth of the workers are women (25%)? | 1. Yes<br>0. No |
| Q35.5 | Would you allow your daughter to work in a workplace where one-tenth of the workers are women (10%)? | 1. Yes<br>0. No |
| Q35.6 | Would you allow your daughter to work in a workplace where there are no women workers (0%)? | 1. Yes<br>0. No |

# S3 Program Implementation and Survey Procedures

## S3.1 Study Setting

The study was conducted in Karnataka, a state in southern India comprising 31 districts that are administratively organized into four divisions: Belagavi, Bengaluru, Kalaburagi, and Mysuru. The study sample consisted of Kasturba Gandhi Balika Vidyalayas (KGBVs) and Karnataka Kasturba Gandhi Balika Vidyalaya Hostels (KKGBVs), residential schools established for girls from socially and economically disadvantaged backgrounds in educationally backward blocks. The study covered all four administrative divisions of Karnataka, except the districts of Kodagu, Dakshina Kannada, and Udupi in the Mysuru Division, where KGBV/KKGBV hostels were not operational during the study period. The official list obtained by the implementation partner from the Karnataka State Education Department for the 2024–2025 academic year contained 218 KGBV/KKGBVs. Of these, 205 schools were selected for inclusion in the study.

Figures S1a and S1b present the geographic distribution of the study sample. Figure S1a shows the locations of participating KGBV/KKGBVs by treatment assignment across Karnataka, while Figure S1b presents the geographic distribution of surveyed parent households interviewed during the endline survey.

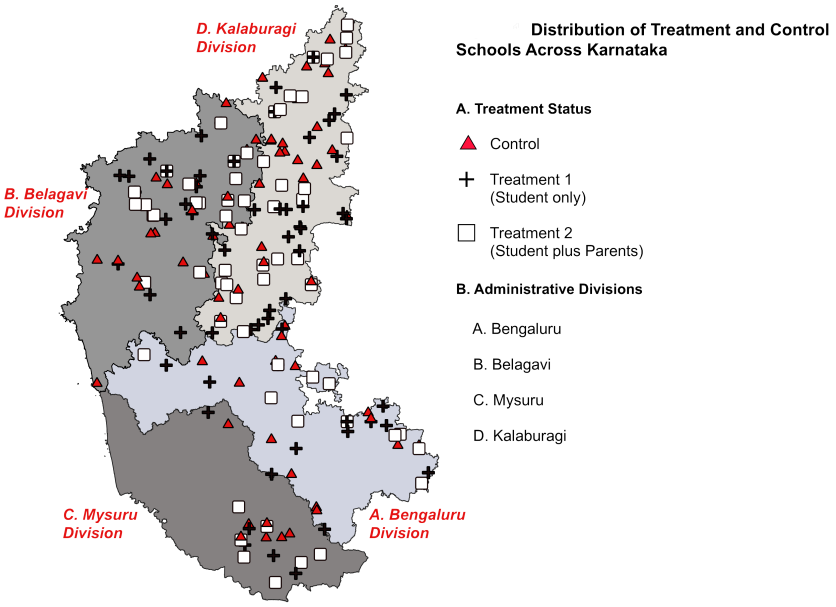
The study targeted girls enrolled in Grades 6 to 9. Four schools were excluded because they were used for pilot testing of the survey instruments prior to the baseline survey. An additional nine schools with fewer than 25 eligible students were also excluded from the study.

The remaining 205 schools were randomly assigned to three study arms comprising two treatment groups and one control group. Randomization allocated 69 schools to Treatment 1, 66 schools to Treatment 2, and 70 schools to the Control group.

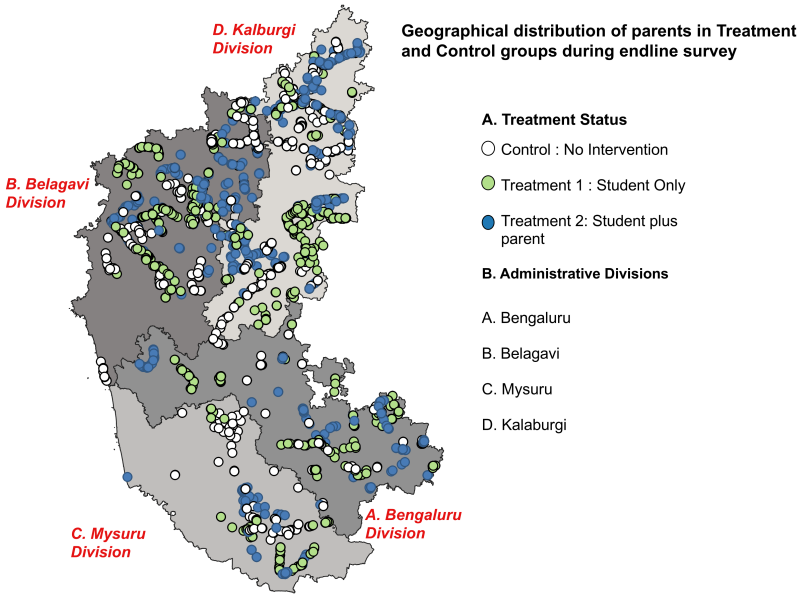
## S3.2 Study Timeline

The entire study timeline spanned August 2024 to April 2026, encompassing instrument piloting, baseline data collection, intervention implementation, and endline surveys. The intervention was implemented over a one-year period, from November 2024 to November 2025. During this period, the student intervention was delivered by hostel wardens through structured English and life-skills sessions in KGBV/KKGBV hostels, while the parent intervention was delivered concurrently through Interactive Voice Response (IVR) calls.

At the beginning of the study period, a pilot survey was conducted in four selected KGBVs during August 2024, followed by the baseline survey in September 2024. Following the




(a) Geographic distribution of participating KGBV/KKGBVs by treatment assignment




(b) Geographic distribution of surveyed parent households at endline

Figure S1: Geographic Distribution of the Study Sample

*Notes:* Panel (a) shows the locations of participating KGBV and KKGBV schools by randomized treatment assignment. Panel (b) shows the geocoded locations of parent households surveyed at endline, classified according to the treatment assignment of the corresponding student's school.

baseline survey, wardens from the treatment schools participated in the first four-day in-person training workshop in October 2024. The student and parent interventions commenced during the second week of November 2024 and continued until the first week of November 2025.

Program implementation was temporarily paused between the last week of March 2025 and May 2025 to coincide with the academic summer break. Before implementation resumed in June 2025, wardens participated in the second four-day training workshop in May 2025. A third four-day warden training workshop was conducted in September 2025 to reinforce program implementation. The parent IVR intervention continued uninterrupted throughout the intervention period.

Program implementation concluded in the first week of November 2025. Following a two-week cooling-off period, the student and warden endline surveys commenced in the fourth week of November 2025, while the parent endline survey began approximately two weeks later and continued through April 2026. Figure 3 presents the overall timeline of the study.

## S3.3 Randomization and Treatment Allocation

This study employed a cluster randomized controlled trial (RCT) design, with the KGBV/KKGBV hostel serving as the unit of randomization. The 205 eligible schools were randomly assigned to one of three study arms.

Randomization was conducted within strata defined by the interaction of the four administrative divisions (Belagavi, Bengaluru, Kalaburagi, and Mysuru) and KGBV school type (Types 1, 4, and 5), resulting in 12 randomization strata. Schools were allocated in approximately equal proportions across the three study arms, resulting in 69 schools in Treatment 1 (T1), 66 schools in Treatment 2 (T2), and 70 schools in the Control group.

The three study arms were defined as follows:

- **Treatment 1 (T1): Student Intervention.** Students received the life-skills curriculum delivered by trained wardens through structured weekly sessions.
- **Treatment 2 (T2): Student and Parent Intervention.** Students received the same life-skills curriculum as those in Treatment 1, while their parents additionally received the Interactive Voice Response (IVR) intervention.
- **Control Group (C): Business-as-Usual.** Schools assigned to the control group continued their regular activities and did not receive either intervention.

Between randomization and the endline survey, three schools experienced administrative changes that affected their inclusion in the final analysis sample. One school assigned to Treatment 1 no longer had students enrolled in the target grades following an administrative decision to transfer students to a neighboring KGBV/KKGBV hostel, which was also assigned to Treatment 1. A second school assigned to Treatment 2 ceased operations due to administrative and operational reasons reported by the Karnataka State Education Department. Finally, one control school was administratively merged with a neighboring Treatment 2 school shortly before the endline survey, although program implementation continued only in the originally assigned treatment schools. Consequently, the final endline sample comprised 202 schools.

## S3.4 Intervention Implementation

The intervention was implemented in partnership with the SwaTaleem Foundation, a non-profit organization that works with adolescent girls, their parents, and other education stakeholders to strengthen foundational English skills, life skills, and parental engagement in girls' education. Within the Karnataka KGBV/KKGBV system, the program adopted a multi-stakeholder implementation model involving students, hostel wardens, and parents. Rather than relying on external facilitators to deliver program activities, the intervention was integrated into the existing hostel ecosystem by positioning hostel wardens as the primary facilitators of the student intervention while simultaneously engaging parents through an Interactive Voice Response (IVR) platform.

## S3.5 Student Intervention

**Implementation Model.** A distinctive feature of the intervention was the decision to deliver program activities through hostel wardens rather than external facilitators, thereby leveraging the hostel ecosystem. Unlike conventional education programs where trainers visit schools periodically, wardens reside within KGBV/KKGBV hostels and interact with students on a daily basis. Their long-term relationship with the girls, familiarity with the hostel environment, and understanding of students' academic and personal needs made them well positioned to facilitate life-skills and English sessions throughout the academic year. Furthermore, this system of curriculum delivery facilitated the successful implementation of the program as an after-school program in complement to the main academic curriculum without compromising the academic hours.

The intervention followed a cascading implementation model. Trainers from the SwaTaleem Foundation first trained hostel wardens through intensive in-person workshops. Following these workshops, wardens received continuous pedagogical support through weekly coaching calls before delivering the English and life-skills sessions to students. This approach allowed

program facilitators to provide continuous guidance while enabling wardens to integrate intervention activities within the regular hostel routine.

**Curriculum Facilitation and Warden Training.** The capacity-building component formed the foundation of program implementation. Prior to the commencement of the intervention, wardens assigned to treatment schools participated in an intensive four-day in-person training workshop. The workshop introduced wardens to the program's objectives and familiarised them with the English and life-skills curricula, lesson sequencing, facilitation techniques, classroom management practices, activity-based learning methods, and approaches for engaging adolescent girls in participatory discussions.

Rather than functioning as a one-time orientation, training was designed as an ongoing process. Two additional four-day in-person training workshops were organized during program implementation. The second workshop was conducted following the academic summer break (May 2025) before program activities resumed for the new academic year, while the third workshop was organized later in the academic year (September 2025) to reinforce lesson delivery and prepare the wardens for the remaining modules.

Between training workshops, the program team facilitators conducted weekly coaching calls with wardens. A total of 25 coaching calls were conducted between November 2024 and November 2025. These calls served multiple purposes. Facilitators reviewed upcoming lesson plans, explained key concepts contained in the English and life-skills modules, discussed anticipated classroom challenges, and provided pedagogical support before each week's sessions. The coaching calls also provided wardens with an opportunity to seek clarification, discuss implementation experiences, and receive feedback on program delivery.

**Curriculum Resources.**To ensure consistency in program delivery, wardens were provided with a comprehensive facilitator handbook containing detailed lesson plans, learning objectives, suggested classroom activities, facilitation guidance, and implementation instructions for every English and life-skills session. Students received corresponding activity workbooks containing exercises and activities aligned with the facilitator handbook.

The intervention adopted an activity-based learning approach rather than traditional lecture-based instruction. Wardens were trained to facilitate sessions using storytelling, role-playing, case studies, games, reflective exercises, and group discussions. The objective was to create an interactive learning environment that encouraged girls to actively participate, discuss personal experiences, and apply concepts through experiential learning rather than passive classroom instruction.

**English Curriculum.** The English component was designed to strengthen foundational English proficiency among students. Lessons followed a structured activity-based approach

that progressively developed vocabulary, listening, speaking, reading, and basic writing skills through interactive classroom activities.

The curriculum was originally intended to provide additional English support for students requiring greater assistance in foundational language skills. However, recognizing the considerable variation in learning levels across hostels, wardens were given flexibility to determine the most appropriate mode of delivery based on the needs of students within their respective hostels. A detailed description of the English curriculum is presented in Appendix (Table S1)

**Life-skills Curriculum.** The Critical Skills for Adolescent Girls, otherwise, Life Skills curriculum focused on developing socio-emotional competencies among adolescent girls. Modules addressed topics including communication skills, self-awareness, emotional regulation, problem-solving, decision-making, mental health, personal hygiene, balanced nutrition, and aspirations etc.

Consistent with the overall pedagogical philosophy of the program, lessons emphasized active participation rather than rote learning. Sessions encouraged discussion, reflection, collaborative activities, and practical application through storytelling, role plays, and case-based exercises. A detailed description of the curriculum is presented in Appendix(Table S2).

**Session Delivery and Operational Flexibility** Taking into account the number of weeks coaching calls were scheduled, the curriculum aimed at delivering 25 weeks of English and life-skills lessons. The intervention was designed to deliver two 40-minute English sessions and one 40-minute life-skills session each week, thus targeting a 50-hour intervention program for students. However, recognizing the substantial variation across KGBV types, wardens were provided flexibility in organizing session schedules while maintaining the intended curriculum coverage.

For example, In Type 1 KGBVs (grade 6, 7, 8), where schools and hostels operate together and student enrollment often exceeds 150 girls across multiple grades, wardens frequently divided students into smaller groups and conducted multiple sessions according to grade level. This approach enabled more effective classroom management while ensuring that all students received the intended curriculum.

In contrast, Type 4 and Type 5 KGBVs and KKGBVs, which function primarily as residential hostels with comparatively smaller enrollment, joint sessions were conducted for all eligible students. This system allowed wardens to adapt delivery according to hostel context improved programme feasibility while preserving the overall structure and learning objectives of the intervention.

**Implementation Challenges and Programme Adaptation.**Programme implementation required continuous adaptation to operational realities within the KGBV system. One

recurring challenge was the transfer and replacement of hostel wardens during the academic year. Because wardens are periodically reassigned for administrative reasons, newly appointed wardens occasionally joined treatment schools after programme implementation had already commenced.

To minimise disruption, the programme team closely monitored staff movements throughout the intervention period. Newly appointed wardens were rapidly identified and integrated into the programme through orientation sessions, coaching calls, and access to facilitator materials, enabling them to continue lesson delivery with minimal interruption.

Another implementation consideration concerned the possibility of spillovers between treatment and control schools. In several locations, multiple KGBVs operated within the same campus or in close proximity while belonging to different treatment arms. To minimise contamination, only wardens assigned to treatment schools participated in programme training workshops and coaching calls. Programme staff carefully verified participant lists before each training session and repeatedly instructed wardens not to share intervention materials or conduct programme activities outside their assigned schools.

Although wardens maintained attendance registers for programme activities, the demanding nature of their administrative and caregiving responsibilities limited the availability of complete session-level attendance records of students across all hostels. Consequently, programme monitoring relied on multiple complementary sources of information, including coaching-call participation, classroom observations, workshop attendance, and regular communication between facilitators and wardens.

## S3.6 Parent level Intervention

The parent component of the intervention was implemented among parents of students assigned to Treatment 2 using an Interactive Voice Response (IVR) platform. The intervention delivered pre-recorded audio messages directly to parents' mobile phones with the objective of strengthening parental engagement in their daughters' education, improving communication between parents and daughters, and promoting supportive educational and gender-related attitudes. Unlike conventional awareness campaigns that rely primarily on one-way dissemination of information, the IVR intervention adopted a behavioural communication approach using storytelling, practical examples, and behavioural nudges to encourage reflection and sustained engagement among parents.

**IVR Content Developemnt.** The IVR messages, delivered under the programme title "Kishori Vani", were developed by the design and content development team of the SwaTaleem Foundation and contextualised for the Karnataka setting. The research team supported the content development process by synthesising evidence from the literature and

identifying key behavioural barriers related to girls' education, gender norms, parental aspirations, social norms, and parent-child communication.

The messages covered a broad range of themes, including the importance of girls' education, educational and career aspirations, communication between parents and daughters, adolescent health and well-being, gender attitudes, parental engagement with hostel life, and support for girls' educational transitions. A detailed week-wise description of the IVR curriculum is provided in Appendix Table S3

To maintain listener engagement, the messages were developed in two formats:

- **Nudges Only** Informational nudges, providing concise evidence-based messages and practical guidance;
- **Nudges and Prompts** Interactive messages, combining informational or narrative content with simple response options that encouraged parents to actively engage with the programme.

Each Kishori Vani message was limited to approximately two minutes in duration. The relatively short message length was intended to minimise respondent burden while encouraging regular participation throughout the intervention period. Every call began with a brief introduction explaining the purpose of the message before transitioning into the main content. Rather than relying on lecture-style information delivery, the intervention emphasised relatable stories, practical situations, and behavioural nudges to communicate programme messages.

**IVR Delivery and timeline.** The IVR intervention commenced during the second week of November 2024. During the first three weeks, parents received introductory calls designed to familiarise them with the IVR platform, explain the objectives of the programme, and encourage continued participation before the main intervention content began.

Following the introductory phase, one IVR message was delivered each week over a planned implementation period of 52 weeks, with the final message delivered during the first week of November 2025. Unlike the student intervention, which was temporarily paused during the academic summer break, the IVR intervention continued uninterrupted throughout the implementation period.

Calls were typically scheduled on Sunday evenings between 7:00 pm and 08:00 pm, a time identified as convenient by most parents. For each weekly message, the IVR platform made up to four call attempts if the initial attempt was unsuccessful. Programme implementation was continuously reviewed using administrative call records. Beginning in January 2025, an

additional repeat-call system was introduced whereby parents who listened to less than 50 percent of a weekly message received follow-up calls on Monday morning and again on Friday evening. Following an internal review of programme performance, the repeat-call strategy was subsequently discontinued.Similarly, an effort was made during the implementation period to personalise the IVR messages by incorporating the student's name into the audio messages. However, based on feedback received from parents during programme monitoring in April 2025, this feature was discontinued.

**Parent IVRS Database.** Parent contact information was compiled by the programme team in collaboration with hostel wardens prior to programme implementation. At the start of implementation, the programme targeted approximately 4,277 parent phone numbers corresponding to students enrolled in Treatment 2 schools. Because multiple students from the same household could share a mobile phone, programme delivery and monitoring were conducted primarily at the phone-number level rather than the individual parent level.

The contact database remained largely unchanged during the first part of the intervention, with a major update undertaken in August 2025. This update was conducted primarily to accommodate parent requests and preferences regarding contact numbers and to incorporate approximately 868 students newly admitted in 2025. Consequently, parents whose contact information appears only in the 2024 contact list had already been exposed to more than half of the planned intervention period before the database was updated. Following the August 2025 update, the active programme contact list included approximately 4,500 parent phone numbers.

Contact records were subsequently cleaned and reconciled across the two versions of the database. The resulting consolidated database contained 6,555 unique contact records: 2,184 (33.3%) appeared only in the 2024 records, 2,189 (33.4%) only in the 2025 records, 1,827 (27.9%) retained the same phone number across both versions, and 355 (5.4%) corresponded to the same student across both versions but with a different phone number. These differences therefore reflect the programme's mid-implementation contact-list update—including new admissions and requested changes to contact information—rather than a complete replacement of the parent population between implementation periods.

**Monitoring Parent engagement.** The IVR platform automatically generated detailed administrative records for every call delivered during the intervention period. These records included information on whether a call was answered, the duration of listening, and the proportion of the message completed.

For programme monitoring, calls were classified into three categories:

- **Answered calls,** defined as calls with a listening duration greater than zero seconds;

- **Failed calls,** including invalid numbers, unanswered calls, or busy lines; and
- **Successful calls,** defined as calls in which the parent listened to at least 50 percent of the audio message.

For the purposes of this study, programme engagement is measured using the final definition of successful calls, as listening to at least half of the message was considered to represent meaningful exposure to the intervention. Accordingly, a successful IVR week is defined as a week in which a parent listened to 50 percent or more of the scheduled IVR message. This measure forms the primary indicator of parental engagement used throughout the analysis.

# S4 Data Collection.

Data were collected at multiple points before, during, and after programme implementation. The research team used a combination of student, parent, and warden surveys; classroom observations; intermediate monitoring surveys; and programme administrative records. Survey instruments and field protocols were piloted prior to the major rounds of data collection, and data-quality procedures were incorporated throughout fieldwork and subsequent data processing. Table A1 reports the baseline, endline, and matched samples for students, parents, wardens, and KGBVs by treatment arm.

## S4.1 Baseline Survey

**Baseline Survey.** Prior to the baseline survey, the research team conducted preliminary field visits to four KGBVs in the study area to pilot the survey instruments and field procedures.

**Enumerator Training.**Before baseline data collection commenced, enumerators participated in a comprehensive training programme covering survey protocols, ethical considerations, informed consent procedures, respondent identification and tracking, use of SurveyCTO, and administration of paper-based questionnaires. Mock interviews and practice sessions were conducted prior to field deployment.

**Student Baseline Survey.** The student baseline survey was conducted between 15 September and 30 September 2024. The target student population comprised students enrolled in Grades 6–9 in the 2024–25 academic year in KGBVs randomised into the trial. Two of the 205 randomised KGBVs were not operational at the time of baseline data collection and could not be surveyed. Baseline sample sizes and their distribution across treatment arms are reported in Table A1.

Student questionnaires were administered using paper-based survey booklets together with

Optical Mark Recognition (OMR) answer sheets. Each student was assigned a unique identifier to facilitate respondent tracking and longitudinal linkage between baseline and endline records. Student demographic and identifying information required for tracking was recorded separately and digitised through manual data entry.

Prior to participation, written parental or guardian consent was obtained through consent forms distributed to students.

**Warden Baseline Survey.** The warden baseline survey was conducted concurrently with the student baseline survey and was administered electronically using SurveyCTO. The questionnaire incorporated built-in validation checks, skip patterns, and consistency constraints to minimise data-entry errors during fieldwork.

Baseline and matched warden sample sizes are reported in Table A1.

**Parent Baseline Survey.** The parent baseline survey was administered using SurveyCTO between October and November 2024. Parents were randomly selected across the treatment and control groups using an on-the-spot lottery procedure. Electronic questionnaires incorporated validation rules, skip patterns, and consistency checks to improve data quality. The parent baseline sample and its distribution across treatment arms are reported in Table A1.

The parent component of the intervention commenced only after completion of the parent baseline survey.

## S4.2 During Intervention Survey.

Several monitoring and intermediate data-collection activities were undertaken throughout the intervention period (August 2024–November 2025) to assess implementation fidelity, participant engagement, programme delivery, and intermediate outcomes. These activities included workshop surveys, a specialised warden survey experiment, classroom observations, a time-use survey, and programme attendance records.

**Warden Workshop Surveys.** To assess changes associated with the capacity-building workshops, pre- and post-training surveys were administered during each of the two rounds of warden workshops. The surveys collected information on wardens' knowledge, perceptions, confidence, and feedback relating to the workshop content and its implementation.

During the first workshop round, 98 wardens completed the pre-training survey and 111 completed the post-training survey. During the second workshop round, 108 wardens completed the pre-training survey and 109 completed the post-training survey. Differences between the pre- and post-training sample sizes reflect participation and survey availability at the respective data-collection points.

The surveys were administered immediately before and after the relevant training activities using structured questionnaires. Responses were used to assess immediate changes in outcomes associated with the training and to inform subsequent workshop delivery and programme refinement.

**Warden Survey Experiment.** A specialised survey was administered to wardens immediately following the first round of in-person warden training. The survey was designed to assess English proficiency together with socio-emotional and behavioural characteristics and included a behavioural game involving monetary incentives. A total of 200 wardens participated in the survey experiment.

Participation in the behavioural component was entirely voluntary. Wardens were informed about the purpose of the exercise, payment procedures, and their right to decline participation at any stage. The questionnaire component was administered using SurveyCTO, whereas the behavioural game was conducted using paper-based materials to facilitate implementation in the field. Prior to data collection, enumerators participated in a dedicated two-day training covering study protocols, informed consent, behavioural-experiment procedures, administration of the different components of the survey, and ethical procedures relating to incentive payments.

**Classroom Observation Surveys.** Two rounds of classroom observations were conducted to monitor implementation fidelity and assess wardens' classroom practices:

Round 1: February–March 2025 Round 2: July–September 2025

The analytical classroom-observation data contain observations from 124 KGBVs in Round 1 and 98 KGBVs in Round 2. The smaller number of observations in Round 2 primarily reflects scheduled sessions that were cancelled or did not take place on the day of the observation visit.

Observation schedules were coordinated with the implementation team to identify teaching sessions while ensuring that wardens were not informed in advance of the specific dates on which their sessions would be observed. This approach was intended to reduce the likelihood of changes in teaching behaviour arising from advance notice of an observation. Observations were conducted using structured paper-based observation schedules, given the complexity of the observation instrument and the requirements of real-time classroom coding.

Before each observation round, enumerators received a dedicated two-day training focused on observation protocols, consistent application of the observation instrument, and inter-enumerator consistency.

**Time-Use Survey.** A paper-based time-use survey was conducted to document wardens'

allocation of time across teaching, administrative responsibilities, and other school-related activities. A total of 198 wardens completed the time-use survey.

**Attendance Records.** Programme attendance records were maintained throughout the intervention by the implementation team. These records captured wardens' participation in the in-person training workshops and coaching-call cycles. The administrative attendance records were subsequently verified, cleaned, and validated by the research team prior to analysis.

## S4.3 Endline Survey

The endline surveys were conducted for three stakeholders; Students, Wardens and Parents. In addition to this, survey was also conducted for Siblings of the students during the parent survey.

**Piloting of Endline Instruments.** Prior to the endline survey, all survey instruments underwent extensive piloting in four KGBVs. The pilot survey assessed the feasibility, timing, respondent comprehension, survey flow, and field implementation of both modules retained from the baseline survey and components newly introduced at endline.

An important addition at endline was an incentivised cognitive and behavioural assessment module administered at the conclusion of the student survey. The module comprised three activities: an English assessment, a Raven's Progressive Matrices assessment, and a task designed to measure public and private confidence.

To ensure that the English assessment was appropriate for students of different ages and grade levels, two versions with different levels of difficulty were developed: one for students in Grades 6–8 and another for students in Grades 9–10. Based on findings from the pilot, a single version of the Raven's assessment and the public- and private-confidence task was administered across grades.

The pilot survey examined students' comprehension of the incentivised activities, the appropriateness and practicality of providing non-monetary incentives, the clarity of task instructions, field logistics, and overall survey flow. Feedback from the pilot informed revisions to questionnaire wording, task instructions, sequencing of survey modules, and implementation protocols before the commencement of full-scale endline data collection.

**Enumerator Training.** Prior to the endline survey, enumerators completed a five-day training programme covering survey administration, respondent identification and tracking, informed consent procedures, quality-assurance protocols, administration of the newly introduced cognitive assessments and incentivised tasks, and field implementation procedures.

Training included extensive field demonstrations, practice interviews, and mock survey sessions prior to deployment. Particular attention was given to the standardised administration of the cognitive and incentivised activities to ensure consistency across enumerators and study locations.

**Student and Warden endline surveys.** The student and warden endline surveys were conducted between 24 November and 26 December 2025.

During the intervention period, one control KGBV merged with a nearby treatment KGBV. Consequently, the number of study KGBVs at endline differed from that at baseline. The baseline, endline, and matched samples by treatment arm are reported in Table A1.

As at baseline, the student endline survey was administered using paper-based questionnaires together with Optical Mark Recognition (OMR) answer sheets. Students enrolled at baseline had advanced by one grade by the time of the endline survey, such that the original study cohort was in Grades 7–10 at endline. The endline survey additionally covered newly admitted students in the relevant grades.

Unique student identifiers established through the survey tracking system were used to identify baseline participants and facilitate longitudinal linkage between baseline and endline records. Student demographic information was verified and updated where necessary during fieldwork. The baseline, endline, and successfully matched student samples are reported in Table A1.

The warden endline survey was administered using SurveyCTO, with built-in validation checks, skip patterns, and consistency constraints incorporated into the electronic questionnaire to improve data quality. Changes in wardens occurred in some KGBVs over the study period; the baseline, endline, and matched warden samples are reported in Table A1.

Fieldwork was organised to minimise disruption to school activities. Surveys were administered during appropriate free periods and with prior approval from school authorities to ensure that students' instructional time and daily routines were minimally affected. Enumerator allocation across schools was determined partly according to student enrolment to facilitate timely completion of fieldwork while maintaining data quality.

**Parent Endline Survey.** The parent endline survey commenced in late November 2025 and continued until April 2026, reflecting the additional time required to locate and track respondents across the study area.

The survey was administered electronically using SurveyCTO, with built-in validation rules, skip patterns, and consistency checks incorporated into the questionnaire. Multiple respondent-tracking strategies were employed to maximise follow-up while ensuring informed consent,

confidentiality, and accurate linkage with baseline records. Baseline and endline parent sample sizes and their distribution across treatment arms are reported in Table A1.

Because parent contact information could change during the intervention period, available identifiers and programme contact records were used during data processing to reconcile updated contact information and facilitate linkage across data sources.

**Sibling Survey.** As part of the parent endline fieldwork, an additional sibling survey was administered to eligible siblings of study participants. A total of 1287 siblings of baseline students were survey during parent survey.

The sibling survey consisted of two components. The first compoenent comprised a short questionnaire containing 18 survey questions collecting information on educational aspirations, general awareness, and related socio-economic characteristics. The second consisted of two cognitive assessments administered to participating siblings.

Both components of the sibling survey were administered using paper-based instruments to facilitate implementation in the field. Interviews were conducted only after obtaining parental or guardian consent, and younger respondents were interviewed in the presence of a parent or guardian where appropriate.

## S4.4 REVIEW NEEDED- Wardens as Key Stakeholders

Wardens play a central role in the day-to-day functioning of KGBV hostels and were key stakeholders in the implementation of the intervention. In addition to their administrative responsibilities, wardens oversee students' academic and extracurricular activities, provide pastoral and socio-emotional support, and serve as an important link between students, schools, and parents. Their role was particularly relevant to the student intervention, as wardens facilitated the delivery of the structured Life Skills and Foundational English lessons. We therefore complement the main student-level results with evidence on wardens' characteristics, responsibilities, time allocation, experiences with programme delivery, and selected attitudes and socio-emotional outcomes.

Information on wardens was collected at several points over the course of the study. The baseline and endline surveys contain information on wardens' characteristics, attitudes, responsibilities, and allocation of time across activities. A dedicated time-use survey conducted during the intervention provides a more detailed account of wardens' daily responsibilities. Programme attendance and classroom- observation data provide direct measures of wardens' engagement with intervention activities; these measures are discussed separately as part of the implementation-fidelity analysis in Section 4. Unless otherwise indicated, the analysis in this section is intended to provide complementary evidence on wardens as key programme

stakeholders rather than an additional family of primary study outcomes.

## S4.5 Warden Characteristics

At baseline, wardens were, on average, approximately 35 years old and had around six years of experience in the role. Approximately one-third to two-fifths held a degree in education, while prior formal training varied across study arms. Appendix Table A2 reports balance across baseline warden characteristics. Although some individual characteristics, including prior formal training and social norms, vary across study arms, joint tests provide no evidence of systematic baseline imbalance. Wardens' employment arrangements also varied, including full-time, part-time, and agency-based positions.

### S4.5.1 Warden Continuity

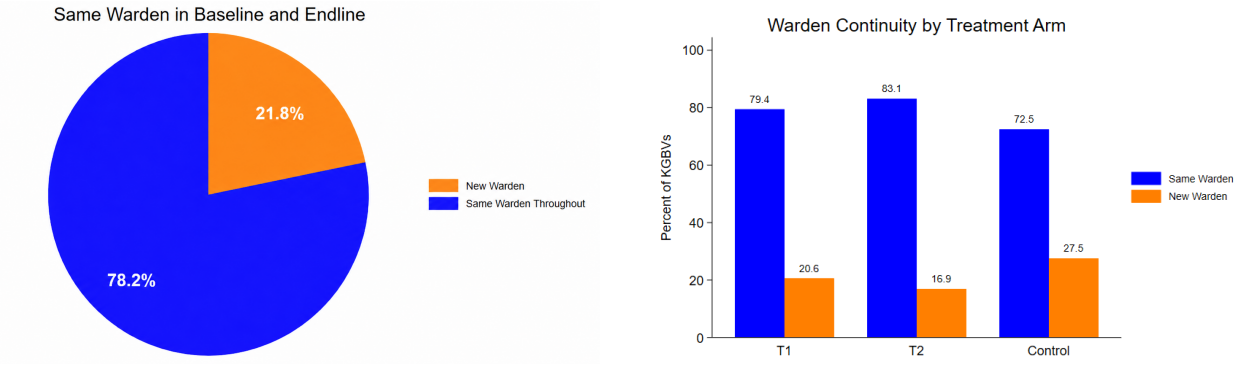
Warden turnover over the study period is an important consideration for interpreting the warden-level evidence. Figure S2 describes warden continuity between baseline and endline. Overall, 78.2 percent of KGBVs retained the same warden, while 21.8 percent had a different warden at endline.

Continuity was relatively high across all three study arms, although there was some variation across groups. The same warden remained in place in 79.4 percent of T1 KGBVs and 83.1 percent of T2 KGBVs, compared with 72.5 percent of control KGBVs. Correspondingly, the proportion with a new warden at endline was 20.6 percent in T1, 16.9 percent in T2, and 27.5 percent in the control group. These differences are presented descriptively and are not interpreted as treatment effects.

Warden continuity is particularly relevant because wardens were directly involved in programme implementation, including training and the facilitation of intervention activities. At the same time, approximately one-fifth of KGBVs experienced a change in warden between baseline and endline. Consequently, baseline and endline warden measures do not always refer to the same individual. We therefore interpret comparisons of warden outcomes across survey rounds as comparisons of wardens serving in the sampled KGBVs rather than as individual-level longitudinal changes, unless the analysis explicitly restricts the sample to wardens observed at both baseline and endline.

### S4.5.2 Warden Responsibilities and Time Use

Wardens combine a wide range of responsibilities within KGBV hostels, including academic support, administration, student wellbeing, pastoral support, and extracurricular activities. We examine how wardens allocate their time across these responsibilities using information collected at baseline, during the dedicated time-use survey, and at endline. Table S19 reports




(a) Overall warden continuity

(b) Warden continuity by treatment arm

Figure S2: Warden Continuity Between Baseline and Endline

*Notes:* Panel A shows the overall proportion of KGBVs retaining the same warden between baseline and endline. Panel B reports warden continuity separately by treatment arm. Overall, 78.2 percent of KGBVs retained the same warden, while 21.8 percent had a new warden at endline. The corresponding shares retaining the same warden were 79.4 percent in T1, 83.1 percent in T2, and 72.5 percent in the control group.

mean hours by experimental group. Because the three survey instruments differ in the level of detail with which time use was elicited, we focus primarily on comparisons across experimental groups within each survey round rather than changes in reported hours across rounds.

At baseline, patterns of time allocation were broadly similar across the three experimental groups. Wardens reported approximately 2.8–3.2 hours of academic responsibilities, 2.5–2.8 hours of administrative responsibilities, and 2.1–2.3 hours of other responsibilities. The dedicated time-use survey, which elicited activities in considerably greater detail during programme implementation, shows more pronounced differences in reported academic time. On weekdays, T1 wardens reported an average of 6.5 hours of academic activity, compared with 5.7 hours among T2 wardens and 5.2 hours among control wardens.

Part of this difference reflects time devoted to ST, an intervention-specific academic activity. Excluding ST, reported weekday academic time falls to 5.3 hours in T1 and 4.5 hours in T2, compared with 4.9 hours in the control group. The corresponding weekend estimates excluding ST are 5.5 hours in T1, 4.9 hours in T2, and 6.6 hours in control schools. Thus, the relatively high academic time observed for T1 during implementation does not appear to reflect a general increase in non-programme academic responsibilities. Instead, an important part of the difference is associated with programme-specific activity.

By endline, reported time allocation is again broadly similar across the experimental groups. Weekday academic time ranges from 2.5 to 2.8 hours, administrative time from 1.9 to 2.0 hours, and other responsibilities from 5.4 to 5.8 hours. A similar pattern is observed for

weekend responsibilities. Taken together, these descriptive patterns suggest that the additional academic time observed during programme implementation was concentrated in programme-related responsibilities rather than reflecting a persistent reallocation of wardens' time toward academic activities.

Table S19: Warden Time Allocation Across Survey Rounds

| | **T1** | **T2** | **Control** |
|---|---|---|---|
| *Panel A: Baseline* | | | |
| Academic responsibilities | 3.19 | 3.09 | 2.83 |
| Administrative responsibilities | 2.79 | 2.78 | 2.53 |
| Other responsibilities | 2.15 | 2.16 | 2.32 |
| *Panel B: Dedicated Time-Use Survey — Weekday* | | | |
| Academic responsibilities | 6.50 | 5.66 | 5.19 |
| Academic responsibilities excluding ST | 5.29 | 4.55 | 4.94 |
| Administrative responsibilities | 2.12 | 1.89 | 2.40 |
| Other responsibilities | 7.30 | 5.77 | 7.77 |
| *Panel C: Dedicated Time-Use Survey — Weekend* | | | |
| Academic responsibilities | 6.88 | 6.37 | 6.87 |
| Academic responsibilities excluding ST | 5.53 | 4.94 | 6.56 |
| Administrative responsibilities | 2.46 | 2.45 | 2.79 |
| Other responsibilities | 7.90 | 7.08 | 8.06 |
| *Panel D: Endline — Weekday* | | | |
| Academic responsibilities | 2.55 | 2.57 | 2.78 |
| Administrative responsibilities | 2.00 | 1.86 | 1.89 |
| Other responsibilities | 5.42 | 5.77 | 5.72 |
| *Panel E: Endline — Weekend* | | | |
| Academic responsibilities | 2.56 | 2.91 | 3.12 |
| Administrative responsibilities | 2.13 | 2.05 | 2.03 |
| Other responsibilities | 6.35 | 6.46 | 6.54 |

*Notes:* The table reports mean hours allocated to academic, administrative, and other responsibilities by experimental group. "Other responsibilities" combine student wellbeing, pastoral, co-curricular, and other warden duties. In the dedicated time-use survey, academic responsibilities include primary academic duties and additional academic activities. ST is an intervention-specific academic activity; Panels B and C therefore additionally report academic time excluding ST. At endline, wardens who reported that their weekend responsibilities were the same as their weekday responsibilities are assigned their corresponding weekday hours when constructing the weekend measures. Where the sum of reported activities exceeded 24 hours, academic and administrative time were retained and the residual "other" category was adjusted so that total reported time did not exceed 24 hours. Sample sizes vary across outcomes because of item-level nonresponse. Because the baseline, dedicated time-use, and endline instruments differ in the level of detail with which activities were elicited, reported levels should not be interpreted as directly comparable across survey rounds. All comparisons are descriptive.

### S4.5.3 Wardens' Attitudes, Socio-Emotional Outcomes, and Behaviour

We next examine whether the intervention was associated with selected dimensions of wardens' attitudes and socio-emotional outcomes. We consider standardized indices of gender

attitudes, perceived social norms, self-esteem, perceptions of students' abilities, and stress. Table **??** reports estimates separately for T1 and T2 relative to the control group. Regressions include baseline warden characteristics and randomization-stratum fixed effects, with standard errors clustered at the KGBV level.

**Warden-reported outcomes.** Overall, we find limited evidence of differences in wardens' attitudes and socio-emotional outcomes across treatment arms. Estimated effects on gender attitudes are positive for both T1 and T2 relative to the control group (0.379 and 0.323, respectively), but neither estimate is statistically significant. Similarly, we find no statistically significant effects on perceived social norms or self-esteem. Wardens in both treatment arms also report somewhat more positive perceptions of students' abilities than wardens in the control group, with coefficients of 0.405 for T1 and 0.297 for T2, although these differences are not statistically significant.

Stress is the one outcome for which we observe suggestive evidence of a difference. The estimated effect for T2 is positive (0.390; $p = 0.056$), indicating higher reported stress relative to the control group, although the estimate falls just short of conventional statistical significance at the 5 percent level. The corresponding T1 estimate is smaller (0.189) and statistically insignificant. Moreover, the difference between the T1 and T2 coefficients is not statistically significant. Taken together, these results provide little evidence of systematic effects on wardens' attitudes or socio-emotional outcomes. The suggestive increase in stress among T2 wardens is therefore interpreted cautiously.

**Parent-reported changes in warden behaviour.** We complement wardens' self-reported outcomes with parents' reports of changes in warden behaviour. Table S20 examines whether parents reported no change in warden behaviour, increased teaching, or increased caring. These outcomes provide a complementary perspective on whether changes in wardens' engagement were visible to students' families.

For T1, there is little evidence of a change in parent-reported warden behaviour. The estimated effects on reporting no change (0.1 percentage points), increased teaching (8.9 percentage points), and increased caring ($-4.8$ percentage points) are all statistically insignificant. In contrast, parents in T2 are 12.8 percentage points less likely to report no change in warden behaviour relative to the control group ($p < 0.05$) and 13.7 percentage points more likely to report increased teaching ($p < 0.05$). The estimated effect on increased caring is positive at 6.0 percentage points but is not statistically significant.

Direct comparisons between the two treatment arms provide some additional evidence that these patterns are stronger under T2. Parents in T2 are 12.9 percentage points less likely

than those in T1 to report no change in warden behaviour ($p = 0.036$). The difference in increased teaching between T2 and T1 is smaller and statistically insignificant (4.9 percentage points; $p = 0.450$). The T2–T1 difference in increased caring is 10.8 percentage points and is marginally significant ($p = 0.053$).

Taken together, the parent reports provide evidence that changes in warden behaviour were more visible under T2, particularly through increased teaching relative to the control group. This evidence complements the warden-reported measures above but captures a distinct dimension of programme implementation: parents' perceptions of changes in wardens' behaviour rather than wardens' own assessments of their attitudes and socio-emotional outcomes.

Table S20: Effects on Parent-Reported Changes in Warden Behaviour

| | **Parent-reported change in warden behaviour** | | |
|---|---|---|---|
| | **No change** | **Teaching increased** | **Caring increased** |
| | (1) | (2) | (3) |
| T1 (Lesson) | 0.001 | 0.089 | -0.048 |
| | (0.058) | (0.057) | (0.055) |
| T2 (Lesson + IVR) | -0.128** | 0.137** | 0.060 |
| | (0.060) | (0.060) | (0.057) |
| Control mean | 0.421 | 0.401 | 0.396 |
| Observations | 4,216 | 4,216 | 4,216 |
| KGBV clusters | 203 | 203 | 203 |
| Parent controls | Yes | Yes | Yes |
| KGBV controls | Yes | Yes | Yes |
| Endline strata FE | Yes | Yes | Yes |
| T2 − T1 | -0.129** | 0.049 | 0.108* |
| $p$-value: T2 = T1 | 0.036 | 0.450 | 0.053 |

*Notes:* The table reports treatment-effect estimates for parent-reported changes in warden behaviour. T1 denotes the lesson-only intervention and T2 denotes the lesson plus IVR intervention. The omitted treatment category is the control group. All specifications include parent controls, KGBV controls, and endline randomization-stratum fixed effects. Standard errors clustered at the KGBV level are reported in parentheses. The row "T2 − T1" reports the difference between the two treatment coefficients, with the corresponding $p$-value reported below. Control-group means are reported for reference. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

### S4.5.4 Wardens' Engagement with and Perceptions of the Programme

Wardens were directly involved in implementing the student intervention. As discussed in the implementation-fidelity section, programme attendance and classroom-observation data document wardens' participation in training and programme activities and the extent to which structured lessons were delivered as intended. Rather than reproducing these implementation measures here, we complement them with information on wardens' own experiences of delivering the programme and their perceptions of their responsibilities and

the programme's benefits for students. Figure S3 summarizes wardens' engagement with and perceptions of the programme. Panel A presents self-reported programme delivery, Panel B presents wardens' perceived success in carrying out key responsibilities, and Panel C presents their perceptions of the programme's benefits for students.

**Programme delivery.** Panel A summarizes wardens' self-reported delivery of the five programme modules. Reported implementation was generally high across components, with between 72 and 78 percent of wardens reporting that they delivered at least 80 percent of the intended sessions. These reports are broadly consistent with the implementation-fidelity evidence presented earlier: wardens participated in an average of 17.9 coaching calls during the intervention and were observed to be on task for 83.7 percent and 77.4 percent of class time in the two classroom-observation rounds, respectively. Taken together, these measures indicate substantial, though not complete, warden engagement in programme delivery.

**Perceived success in warden responsibilities.** Panel B examines wardens' assessments of their effectiveness in carrying out key responsibilities associated with their role. Wardens report high levels of perceived success across the different dimensions, including teaching, supporting student wellbeing, classroom management, promoting student engagement, generating positive impacts for students, and fulfilling their overall responsibilities as wardens. More than 90 percent report being successful or very successful in supporting student wellbeing, managing classrooms, and improving student engagement, while nearly 98 percent report being successful or very successful in their overall role. Self-assessed success in teaching English is somewhat lower, although still high. Overall, these responses suggest that wardens generally perceived themselves as effective in carrying out their programme-related responsibilities.

**Perceived benefits for students.** Panel C reports wardens' perceptions of the benefits of the programme for students. Assessments are overwhelmingly positive across the dimensions considered. Almost all wardens agree or strongly agree that the programme supports students' English skills, employment and higher-education opportunities, confidence, ability to overcome challenges, problem-solving skills, and belief in equal opportunities, as well as helping to challenge gender stereotypes. The limited variation across these measures indicates broad warden support for the programme and a strong perception among wardens that it benefits students across both academic and socio-emotional dimensions. These responses should, however, be interpreted as wardens' perceptions of programme benefits rather than as independent evidence of programme impacts.

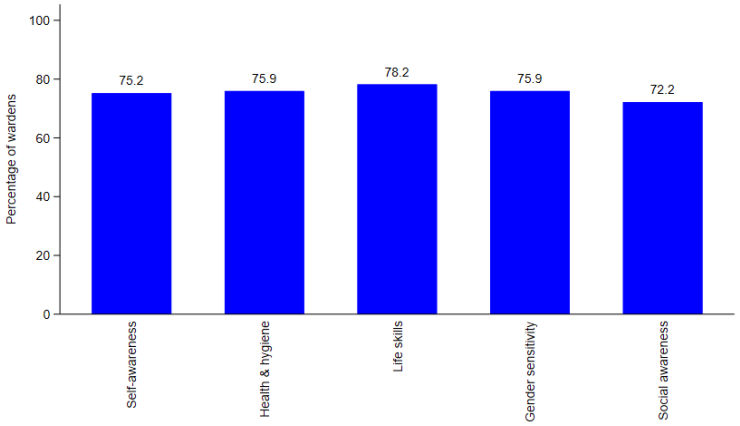
Taken together, the warden-level evidence highlights the breadth of wardens' responsibilities

and their central role in programme implementation. Wardens divide their time across academic, administrative, and broader student-support activities. During implementation, T1 wardens reported relatively high academic time; however, much of the difference relative to control schools is accounted for by programme-specific ST activities. Once ST is excluded, there is little indication of a broader or persistent reallocation of warden time toward academic responsibilities. By endline, reported time allocation is broadly similar across the three experimental groups.

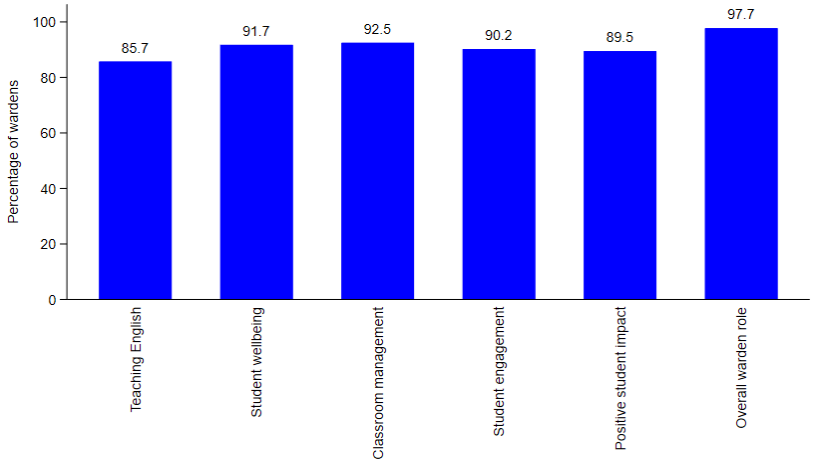
The programme was implemented in a context of relatively high, though incomplete, warden continuity. Approximately 78 percent of KGBVs retained the same warden between baseline and endline, while around 22 percent experienced warden turnover. This turnover is important for interpreting warden-level evidence, since baseline and endline measures do not always correspond to the same individual. At the same time, self-reported programme delivery, coaching participation, and direct classroom observations collectively indicate substantial warden engagement with intervention activities. Wardens also report high perceived success in carrying out their responsibilities and overwhelmingly positive perceptions of the programme's benefits for students.

We find more limited evidence of systematic effects on wardens' own attitudes and socio-emotional outcomes. Estimated effects on gender attitudes, perceived social norms, self-esteem, and perceptions of students' abilities are not statistically significant. T2 wardens report higher stress relative to the control group, although the estimate falls just short of conventional statistical significance at the 5 percent level. Parent reports provide a complementary perspective on warden behaviour: parents in T2 are less likely to report no change in warden behaviour and more likely to report increased teaching relative to the control group, while corresponding effects for T1 are not statistically significant.

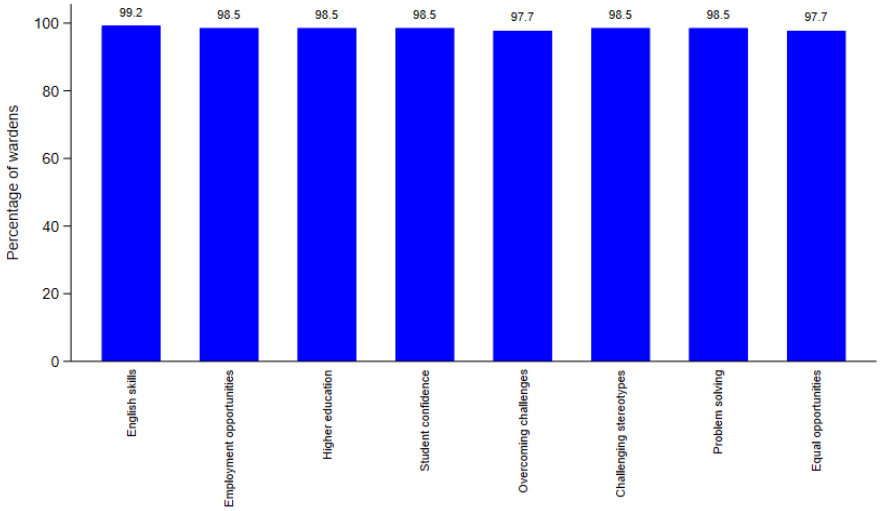
Overall, these findings characterize wardens as active intermediaries between the intervention and students. The evidence on their responsibilities, programme engagement, experiences, and parent-reported behaviour provides useful context for interpreting both the implementation-fidelity evidence and the student-level treatment effects, while the observed warden turnover underscores the importance of distinguishing KGBV-level programme exposure from changes experienced by individual wardens over time.




(a) Programme delivery




(b) Perceived success in warden responsibilities




(c) Perceived benefits for students

Figure S3: Wardens' Engagement with and Perceptions of the Programme

*Notes:* Panel A reports the percentage of wardens reporting delivery of at least 80 percent of the intended sessions for each programme module. Panel B reports the percentage of wardens describing themselves as successful or very successful across selected responsibilities. Panel C reports the percentage of wardens who agree or strongly agree with statements concerning the programme's benefits for students. Measures in Panels B and C are based on wardens' self-reports and should be interpreted as perceptions rather than independent measures of programme impacts.