# Can Economic Opportunities Reshape Gender Norms among Religious Youth? Evidence from Northern Nigeria

Benjamin Crost  Oeindrila Dube  Marcus Holmlund

Eric Mvukiyehe  Emily Crawford*

March 2025

**Abstract**

Can expanding economic opportunities reshape gender norms? We explore this question in Northern Nigeria, a religiously conservative region marked by highly restrictive gender norms. Using a randomized controlled trial, we extend vocational training to impoverished youth, including boys in Quranic schools and girls in Islamic education. The training did not discuss gender equity. Yet, it both increased beneficiaries' incomes, and fostered more equitable views on women's education, labor force participation, and household decision-making. These effects extended beyond beneficiaries to their spouses and parents, highlighting the potential for economic interventions to engender social change in religiously conservative settings.

**Keywords:** Gender Norms, Religion, Labor Market Intervention, Vocational Training

**JEL Classification:** C93, I25, J16, J24, M53, O15

*Crost: University of Calgary. E-mail: benjamin.crost@ucalgary.ca. Dube: University of Chicago and NBER. E-mail: odube@uchicago.edu. Holmlund: World Bank, Development Impact (DIME). E-mail: mholmlund@worldbank.org. Mvukiyehe: Duke University and J-PAL. E-mail: eric.mvukiyehe@duke.edu. Crawford: Binomial Optimus Limited. E-mail: emily.e.crawford@gmail.com.

We are grateful to our partners in the UK's FCDO, in particular to Ed Smithson, Richard Sandall, Chris Price, Ogechukwu Omeribe, Tonia Momodu, Kristen Hopkins, Conor Doyle, and Federica Di Battista. We would also like to express our deep gratitude to the entire Mafita project implementation team, including the former project directors Umar Mohammed, Ana Vinambres and Erling Petersen, and to the broader team at Adam Smith International including Alison Bell and Alex Elphinstone, for their collaboration. We would also like to thank Rachel Glennerster, Madeline McKelway, Sendhil Mullainathan, Emily Nix, Adam Osman, Cyrus Samii, and Srinivasan Vasudevan for insightful comments; as well as participants at the 2021 APSA Annual Meeting and the UC San Diego development seminar for helpful suggestions on earlier versions of the paper. We also thank Aanchal Bagga, Rongmon Deka, Samih Ferrah, Nausheen Khan, Ijeoma David Ukoko and Nuruddeen Umar for critical research assistance and field coordination activities. This study was funded by FCDO Nigeria under the i2i trust fund administered by the World Bank, which contracted endline data collection. It received IRB approval under the National Health Research Ethics Committee of Nigeria (protocol NHREC/01/01/2007-19/10/2018). The study was also pre-registered under Pre-Analysis Plan AEARCTR-0004500.

# 1. Introduction

Gender inequality is a pervasive challenge throughout the world, particularly in developing countries (Duflo, 2012; Jayachandran, 2015). Compared to men, women have lower levels of employment, income, health, education and decision-making power, and experience higher levels of violence. These gender gaps have been remarkably persistent, despite sustained efforts to reduce them. Cultural norms have played a key role in maintaining this inequity (Alesina et al., 2013; Fernández and Fogli, 2009; Fernández, 2013; Bertrand et al., 2015; Bursztyn et al., 2020). For instance, norms that restrict women's personal autonomy and economic opportunities are widespread in many developing countries (Deininger et al., 2013; Field and Vyborny, 2022; Bursztyn et al., 2023; Abou Daher et al., 2023).

In this paper we examine if expanding economic opportunities available to male and female adolescents can help to re-shape inequitable gender norms. Through a field experiment, we study the effects of extending vocational training to youth residing in two states of Northern Nigeria. In the region, religious beliefs and institutions favor a highly conservative interpretation of Islam; and prevailing gender norms limit female autonomy in making decisions related to the labor market. For example, the custom of *kulle* prohibits married women from interacting with non-family members of the opposite sex, restricting women's mobility outside the home (Robson, 2000; Zakaria, 2001; Adeyemi et al., 2016); and fathers typically have the final word on the labor market participation of their unmarried daughters (Csapo, 1981; Erulkar and Bello, 2007; Para-Mallam, 2010; Laouan, 2018; Seff et al., 2022).

The vocational training program we study provides 9 months of trade-specific training. The training takes place in well-equipped workshops and is delivered by teachers who are highly skilled in the trade. Critically, the curriculum does not discuss promoting gender equity or changing gender norms.

Rather, the program aimed to extend economic opportunities to highly impoverished youth,

some of whom are among the most religiously conservative segments of Nigerian society. Our sample of 1,824 youth include boys residing in free Quranic schools under the care of Mallams, girls receiving an Islamic education, orphans, and disabled youth residing in charitable institutions. 900 of these youth were randomized to receive the training, while another 924 served as the control group.

To gauge the economic and social effects of these training opportunities, we measured economic and gender-related outcomes approximately 12 months after the adolescents completed the training. We uncover four sets of results. First, we find that the vocational training substantially improved the economic outcomes of participants, boosting an index of income and employment by .24 standard deviation units (SDs) and an asset index by .18 SDs. Second, we observe that the training also led to similar-sized improvements in gender norms. Trained adolescents displayed .12 SD larger increases in an index of gender norms, compared to those in the control group. In unpacking this result, we observe that trainees were more likely to indicate that women should have a say in decisions related to household lending and borrowing, as well as having children. They were also more likely to support the idea that boys should do as much domestic work as girls, and that girls are as capable as boys in school. The fact that the training did not mention gender or have any gender related objectives limits the concern that these responses reflect social desirability bias, a challenge that the literature on gender and female empowerment generally has to contend with.

Disaggregating the analysis by gender reveals that the effect on girls' gender norms is larger and more precisely estimated. However, we cannot reject the hypothesis that the training affected male and female respondents equally (albeit potentially because we may lack the statistical power to detect small differences in effects across the two sub-samples).

We further find that the training led to a broadening of participants' social interactions. Individuals exposed to the training have a higher percentage of friends who are employed, live outside the respondent's neighborhood, and are of the opposite gender as the respondent.

To explore the relationship between gender norms, employment, and social interactions, we conduct a mediation analysis. This analysis suggests that employment is an important mediator of the training's effect on both social interactions and gender norms. These results are consistent with a mechanism in which trainees' increased employment led to a wider social network, and fostered a change in their gender norms.

Finally, and notably, we also observe that the training led to direct improvements in the gender attitudes of the caregivers of participants, which include parents and husbands. Exposure to the training changes an index of the caregivers' gender norms by 0.16 SDs. This effect is particularly strong for caregivers with a female ward enrolled in the study. Participants' caregivers were more likely to approve of female labor force participation. In addition, they were more likely to indicate that women should occupy leadership positions in society and have a say in decisions about their fertility. These changes in attitude are particularly important in the context of Northern Nigeria where caregivers make critical decisions in areas related to the marriage and employment of their female wards (Csapo, 1981; Erulkar and Bello, 2007; Para-Mallam, 2010; Laouan, 2018; Seff et al., 2022).

The findings in our paper contribute to two distinct literatures. We firstly contribute to the literature on the economic and social effects of job training. Previous research has yielded mixed results on the effectiveness of job training programs in developing countries. Early work in this literature found that vocational training had disappointingly small employment and wage effects (Card et al., 2011; Attanasio et al., 2011; Cho et al., 2013; Adoho et al., 2014; Blattman and Ralston, 2015; Hirshleifer et al., 2016). However, more recent studies have found more promising effects by showing that carefully designed training programs can have transformative effects on the lives of their participants (Attanasio et al., 2017; Field et al., 2019; Alfonsi et al., 2020; Bandiera et al., 2020; Carranza and McKenzie, 2024). Our results complement these studies by showing that job training can generate substantial gains in income and employment even among a highly marginalized population with low labor market attachment in a fragile setting. While most studies in this literature have focused on

the labor market effects of training, a small number of studies has also shown that vocational training programs can have wider social benefits, for example by reducing crime recidivism (Newton et al., 2018) and involvement with armed groups (Blattman and Annan, 2016). Our research adds another critical dimension to the positive social impacts of vocational training by demonstrating its effect on gender norms. These effects highlight the importance of considering broader social benefits in determining the overall effectiveness of job training.

Secondly, we contribute to the literature on gender norms in economics. Several recent experiments show evidence that efforts to re-shape gender norms explicitly can succeed in achieving this objective. This includes interventions that promote school-based discussions of gender equality (Dhar et al., 2022) and life skills training (Edmonds et al., 2023), as well as interventions that combine life skills and vocational training with safe spaces for girls (Bandiera et al., 2020).

Related work has demonstrated the success of economic incentives aimed at achieving ends like delaying early marriage (Buchmann et al., 2023); and shown that assisting women in finding jobs can serve to delay marriage and reduce fertility (Jensen, 2012). Gender empowerment has also been shown to emerge from levers that target cash transfers (Almås et al., 2018), assets (Bandiera et al., 2017; Bedoya et al., 2019; Bossuroy et al., 2022) and independent savings accounts for female beneficiaries (Field et al., 2021).[1].

Our work builds on this literature in two ways. The training we evaluate does not set out to explicitly shift gender attitudes or ameliorate gender inequity by solely promoting female work opportunities. Thus our results show how economic programs that do not target gender norms explicitly, or target women specifically, can also serve to re-shape gender attitudes. Moreover, we show these effects in the context of a society where gender norms are guided by a strict interpretation of Islam, thereby building on a nascent literature that examines how gender attitudes can be changed in such settings. In this space, Bursztyn et al. (2020) ex-

[1]Other gender effects of cash transfers have been shown by Bobonis (2009), Hasan (2010), Attanasio and Lechene (2014), and Armand et al. (2020)

amine how correcting misperceptions of social norms can achieve outcomes like increasing female labor force participation in Saudi Arabia. It is possible that training programs like the one we evaluate can shift gender attitudes in such a context by averting the potential negative signal that traditional gender empowerment programs might send in conservative societies (Buchmann et al., 2023; Ashraf et al., 2020). Indeed, programs perceived to be at odds with traditional norms have the potential to generate backlash in such societies (Glennerster et al., 2021; Blumenstock et al., 2022; Gazeaud et al., 2023).

Third, we document that participant's experience with skills training and improved labor market opportunities have the capacity to influence not only their own, but also their parent's gender views. To our knowledge, this is the first paper to provide evidence for this kind of effect. This finding suggests a broader potential of interventions like this one to change gender norms, particularly in societies where guardians have the final word on female participation in the labor market and female labor supply decisions. Overall, our results are consistent with the existence of a virtuous cycle, in which interventions that improve women's economic opportunities lead to a change in gender norms that in turn creates even more opportunities for women (Duflo, 2012; Jayachandran, 2021).[2] Importantly, our findings point to a concrete policy lever that can be leveraged to kick-start this cycle.

## 2. Institutional Context

In this section, we provide background on poverty and gender norms in Northern Nigeria, two critical aspects of the region's socioeconomic development which are integral to our study.

Our evaluation takes place in the Northern Nigerian states of Kaduna and Katsina, which are both characterized by high levels of poverty. The percentage of people in extreme poverty,

[2]Of course gender norms can also be altered through other conduits besides economic opportunities such as the influx of liberal gender norms from outside (Jensen and Oster, 2009; La Ferrara et al., 2012). And, conversely, female economic outcomes can also be improved through other types of interventions besides those that target gender norms specifically (Ashraf et al., 2020; McKelway, 2018).

under 1 dollar per day (PPP adjusted), is 74 percent in Kaduna and 66 percent in Katsina (Open Data for Africa, 2022). Moreover, unemployment rates are above 25 percent in both states. These economic conditions underscore how extending economic opportunities to impoverished youth in the region could lead to substantial improvements in well-being.

The majority of the population is Muslim in both states. The Nigerian Census has not collected data on religious affiliation since 1963, making it difficult to know the exact percentage of Muslim population by state. However, most estimates place the Muslim population at around 50 percent in Kaduna and above 95 percent in Katsina (Ostien, 2018). Muslims tend to come predominantly from the Hausa-Fulani ethnic group while most Christians are either Igbo or Yoruba. Both states have been under Sharia law since the early 2000s, and moreover, both also adhere to a particularly conservative interpretation of Islam.

## 2.1. Gender norms and institutions in the region's religious setting

Women in this context typically experience low levels of economic and social empowerment. Traditionally, women bear the brunt of domestic work and also exhibit low levels of employment outside the household (Adeyemi et al., 2016; Okojie, 1991; Egbue, 2012; Enfield, 2019). One reason for this low labor force participation is that young women tend to leave school early to marry and have children. In our experimental sample, at age 20, 41 percent of women are married, compared to only 1 percent of men. This lines up with previous estimates of marriage rates in rural Northern Nigeria at age 20 of about 50 percent of women and 4 percent of men (Enfield, 2019). Women also exhibit low levels of participation in civic activities such as politics or community leadership (Egbue, 2012).

The low levels of female empowerment in the region are strongly tied to institutions that prescribe specific gender roles, which are linked to an extremely conservative interpretation of Islam. One of the religious institutions limiting the economic empowerment of women in the region is the practice of seclusion, referred to either by the Hausa term *kulle* or the Arabic *purdah*. In its strictest interpretation, this practice prohibits women from interacting with

non-family members of the opposite sex (Zakaria, 2001). Furthermore, married women are only allowed to leave their house occasionally, accompanied by their husbands (Adeyemi et al., 2016; Robson, 2000). While not all women are subject to this strict interpretation of *kulle*, many of them are restricted in their movement in some way by the practice. For instance, looser interpretations of *kulle* allow married women to leave their houses after dark, for essential visits to kin, or on other occasions with the permission of their husband (Adeyemi et al., 2016; Robson, 2000). Overall, however, these restrictions dramatically limit married women's ability to participate in the labor force and public life more broadly.

While *kulle* traditionally only applies to married women, unmarried women and adolescent girls also tend to have limited decision-making authority in this region. Fathers of unmarried women generally have the final, and often the only, say over decisions, including decisions about education, enrollment in training programs, labor market participation and marriage (Csapo, 1981; Erulkar and Bello, 2007; Para-Mallam, 2010; Laouan, 2018; Seff et al., 2022). For unmarried women whose father is deceased or otherwise unavailable, it is expected that a male relative or other man will step in and act as a guardian who will take these decisions for the young woman in his care.

Other religiously motivated norms limit women's decision-making power within the household. Many women in Northern Nigeria are taught from childhood 'that God decreed that they should remain submissive to their husbands irrespective of their behaviors (Callaway, 1986; Onwutuebe, 2013). This belief has been held to limit their ability to make decisions around fertility and resist intimate partner violence (Adeyemi et al., 2016; Enfield, 2019). Furthermore, inheritance and divorce laws in the region are heavily biased against women, further reducing their bargaining power and status within the household.

The restrictive gender norms of Northern Nigeria are reflected in the attitudes of our sample population. For example, as shown in Appendix Table A1, 88 percent of respondents in the control group said that men should be the sole decision-maker about large household

purchases, and 77 percent said the husband alone should decide whether his wife is allowed to work outside the home. 51 percent of respondents say the husband alone should decide whether the couple has children. 48 percent said that it is more important to educate boys than girls, and only 45 percent said that girls should speak as much as boys in class.

These unequal gender attitudes are widespread even among girls, and among the portion of our study population not enrolled in Islamic education. For instance, 82 percent of girls not in Islamic education say that the husband should be the sole decision-maker for large household purchases, and 65 percent say that the husband should decide the wife's place of work. Even among this group, 30 percent of respondents said that it is more important to educate boys than girls, and only 54 percent said that girls should speak as much as boys in class.[3]

Unequal gender attitudes are also widespread among the participants' caregivers. For instance, as shown in Appendix Table A2, 45 percent of caregivers believe that the husband should be the sole decision-maker about his wife's occupation, and 55 percent believe that the husband alone should decide whether the couple has children.

# 3. Overview of the Training Program

The vocational training program we study, called the Community Skills Development Centers (COSDEC) program, aimed to provide trade-specific training to impoverished youth. The training took place in six different centers, each located in a different part of Kaduna and Katsina state, and was implemented as a part of the Mafita skills-development program funded by the UK Foreign, Commonwealth, and Development Office (FCDO).

The first 3 months of the training consisted of classroom-based instruction in foundational skills (literacy, numeracy, and basic science), as well as "soft skills" training on topics like

[3]These statistics are calculated by restricting the control group in the study to the subset of all other girls except those receiving an Islamic education

goal-setting and interpersonal communication.[4] This was followed by approximately nine months of vocational training in one of seven specific trades: Carpentry and Joinery, Electrical Installation, Fashion Design, Hospitality, Masonry, Office Management, and Welding and Fabrication. This part of the training took place in fully equipped workshops under the supervision of teachers skilled in the trade. Male and female participants were trained together, though the gender composition of the classroom varied substantially across trades and centers. Our analysis in Section 6.5 shows that the effect of the training did not depend significantly on the gender-mix of the classroom.

In addition to the core foundational and vocational training, some COSDEC participants were also offered an additional training in entrepreneurship and business administration that lasted for two months after the end of the vocational training. Our analysis in Section 6.6 also verifies that the effects we observe do not depend critically on this component of the training.

Since participants came from highly impoverished backgrounds, they were paid a stipend of 200 Naira per day (approximately 0.67 USD) for 20 days of the month for the entire duration of their training. This stipend was intended to pay for food and transportation, so that the trainees could focus on their training and did not have to engage in other income-generating activities.

Mafita was designed with great care to be appropriate for local conditions. The curriculum was developed in collaboration with the National Board of Technical Education (NBTE), which oversees vocational education in Nigeria. At the end of their training, participants were encouraged to take the assessment exam and become certified in their trade. Approximately 73 percent of participants completed the program, and 70 percent of participants completed the program and obtained this certification.

[4]The soft skills component made up only a small part of the Mafita training - approximately 45 minutes a week, out of a total of 30 hours of training.

To ensure a high quality of instruction, Mafita also worked with the COSDEC teachers to build their capacities and provide certification through a Master Training Program supervised by the NBTE. Regular classroom and workshop observations were conducted by senior staff to ensure that trainers followed the curriculum and adhered to the NBTE's guidelines for instruction.

Notably, the Mafita program did not explicitly discuss or target gender attitudes in any way. This approach has averted potential backlash from local leaders against what could have been perceived as attempts to impose foreign cultural norms on the local population. For our study, the lack of gender component has the advantage of limiting the potential for social desirability bias among respondents. Since the training did not make gender issues salient, it is unlikely that social desirability would have led respondents who received the training to under-report conservative gender norms relative to respondents who did not receive the training.

# 4. Study Design and Methods

## 4.1. Sample selection and randomization design

Eligibility for the COSDEC training was determined by age, income and membership in one of the marginalized groups targeted by the Mafita program. To be eligible, individuals had to be 15-24 years old and earn less than 226 naira per day (about 0.75 USD at the time of the study). In addition, they had to belong to one of the following marginalized groups: early school leavers, orphans and vulnerable children, persons with disabilities, boys in the "Almajiri" system [5], and girls in Islamiyyah/Quranic education, who are called "IQE" girls.

Since the latter two groups are somewhat specific to the Northern Nigeria context, we discuss them in greater detail here. The Almajiri are boys who have gone through the institution

[5]This includes boys who are currently in an Almajiri school and those who have passed through the Almajiri system.

of Almajiranci, in which poor parents entrust male children to Islamic schools, to study with religious leaders called Mallams. This practice has a long tradition in Northern Nigeria, by some accounts dating back to the 11th century. In the present day, Almajiri tend to have little formal (secular) education and high rates of unemployment. They are largely stigmatized by mainstream society and often perceived to be delinquents and troublemakers.

Girls who receive a religious education receive it in Islamiyyah schools, which focus on Arabic studies and Qur'anic education. In many cases, the Islamiyyah schools do not replace secular schooling, and many girls attend both types of schools. In contrast to the male Almajiri, girls in Islamiyyah schools tend to live at home with their families.

The Mafita program targeted these latter two groups based on widespread concerns that the skills imparted by the Almajiri and IQE education systems have low returns in the labor market, and in the hopes that COSDEC training could expand their labor market opportunities.

To identify eligible individuals, the Mafita project worked with a variety of intake sources that included state agencies such as the State Universal Basic Education Board (SUBEB), State Agency for Mass Education (SAME), as well as a variety of non-governmental organizations (NGOs). Mafita consultants visited intake sources, such as Islamic schools and NGOs that work with marginalized youth, and conducted data collection to ensure that the potential participants met the eligibility criteria.

During the outreach and screening effort, Mafita generated applications from 1824 eligible individuals who applied to their local COSDEC center. Prior to randomization, individuals applied to one of the six centers where the trainings were conducted, specifying in advance the trade they wanted to be trained in. Out of the 1824 eligible applicants, 900 were randomly selected to receive the training, and the other 924 were assigned to the control group. In the randomization, applicants were stratified at the center-trade level since each center had only a limited number of slots available in each trade. Applicants randomized into the

training group were trained from April 2017 to March 2018. And, applicants randomized into the control group were not allowed to apply to a different trade or center for the duration of the study.

## 4.2. Data collection and measurement

Data collection for the endline survey commenced in December 2018 and was completed in May 2019. We developed the survey instrument in collaboration with the program implementers, and refined it through several rounds of pilot testing and field testing. Before data collection began, we developed a pre-analysis plan that specified which outcomes we would measure, as well as the exact survey questions that would be used to measure them. We put particular care into designing the survey instrument to measure the gender norms of youth and their caregivers, two key outcomes of interest for this study. Critically, we measure the gender norms of caregivers to gauge if having a female ward participate in the training alters their gendered perspective. We do this through an extensive array of questions administered to the participants' primary caregivers, most of which are their parents, husbands or Mallams. Like all other pre-specified outcomes, gender norms are measured as an index that aggregates answers to survey questions following the method of Kling et al. (2007).

The survey module on youth gender norms asked an extensive set of questions that covered three broad topics: household decision-making, women's equality and role in society, and domestic violence. In Section 6 below, we report results for the entire index that aggregates 28 questions on gender norms, as well as for sub-indices that separately aggregate the questions for each of the three topics. Control group means and estimated effects of the training on all questions in the gender norms index are reported in Appendix Table A1.

The questions on household decision making asked respondents which spouse in a married couple should be the main decision-maker in a particular household decision, such as making household purchases, lending money, the wife's place of work, etc. The possible answers are the husband alone, the wife alone, or that the decision should be made jointly.

To calculate the index, we assign the value 1 if the respondent thinks the decision should be made by the husband alone, and zero if they think the decision should be made jointly or by the wife alone.

Some of the household decisions we asked about, such as household purchases and lending/borrowing money, are from standard modules (e.g. the DHS: ICF, 2004-2017), while others were adapted to capture the gender norms specific to the local context. For instance, we added questions about whether the wife should work outside the household, her working hours, and her participation in groups, all of which may reflect adherence to the religiously motivated practice of seclusion (*purdah/kulle*).

The questions on domestic violence ask whether a certain behavior by a wife would justify her husband beating her. Examples of behaviors are: if she goes out without telling him, refuses to cook, neglects the children, etc. As in the household decision-making module, some of these questions are standard in modules of this kind while others were designed to reflect the local context (e.g., going out without telling the husband may reflect a violation of seclusion).

The questions about women's equality and role in society were designed to reflect broader norms not directly tied to specific decisions or behaviors. Examples include questions about whether women should occupy leadership positions in society, women should be allowed to work outside the home, and educating girls is as important as educating boys. Some of these questions were previously used by other studies, such as the World Values Survey (Haerpfer et al., 2020), and we developed others for the local context.

The caregiver gender norms outcome is an index that covers three topics: household decision-making, women's equality and role in society, and female labor force participation. The first two topics are based on the same survey questions as the respective topics of the youth gender norms index. The questions about female labor force participation ask whether women of different ages and different marital status should work inside or outside of the household.

For instance, we asked whether it is acceptable that an 18 year old married woman works outside of the household, whether a 22 year old single woman works at home, etc. The questions cover combinations of ages 14, 18 and 22, married and unmarried status, and working inside and outside the home. As for the youth gender norms outcome, we report results for the entire index that aggregates 30 questions on gender norms, as well as for sub-indices that separately aggregate the questions for each of the topics.

We also collected information about participants' social networks by asking them about the characteristics of their five closest friends. While this is not a full social network mapping, it allows us to examine whether the training broadened the composition of participants' social networks, for example by making them closer to individuals of the opposite gender or to individuals from other neighborhoods.

Prior to the start of the training, the Mafita program conducted its own baseline survey on a subset of program participants, as well as on other eligible youth who eventually were not enrolled in the study. This data was collected between October and December 2016 prior to the commencement of the training. Baseline data is available for approximately 85 percent of the study participants. We use data on some variables collected through the baseline survey as additional controls in robustness tests reported in Section 6.6.

# 5. Empirical Strategy

We estimate the effect of the COSDEC training with the following equation:

$$Y_{is} = \beta_0 + \beta_1 T_{is} + X_{is}\Gamma + \alpha_s + \varepsilon_{is} \tag{1}$$

In this equation, $Y_{is}$ denotes the outcome of interest for individual $i$ in randomization strata $s$. The variable $T_{is}$ is an indicator for being randomly assigned to receive the training, and

$X_{is}$ is a vector of pre-specified control variables collected at intake before the randomization was carried out. The equation also includes a set of strata fixed effects, $\alpha_s$. As described in Section 4.1, randomization was stratified at the training-center-by-trade level. As discussed there, participants chose which trade and which center to apply to. We then randomly assigned 30 applicants in each center-trade cell to receive training (this is because there were exactly 30 training slots available for each trade in each center). By controlling for strata fixed effects, we account for the fact that the probability of being randomized into the treatment group was higher for the less popular trade-center combinations. By including these fixed effects, our analysis compares youth who applied for the same trade in the same center, some of whom were randomly chosen to receive training and others to serve as control.

Compliance with the randomization was high – the training was completed by 73 percent of individuals in the treatment group. A small fraction of individuals in the control group (approximately 7 percent) was able to participate in the training. However, the coefficient $\beta_1$ reflects the intent-to-treat effect (ITT) of being randomly assigned to the training group. Estimates of the effect of randomization on training completion are reported in Appendix Table A3, which are large and statistically significant.[6]

Our pre-analysis plan specified 7 primary outcomes, all of which are measured by mean effects indices following Kling et al. (2007).[7] As pre-specified in our analysis plan, for these outcomes, we include estimates that adjust for multiple hypothesis testing by controlling the False Discovery Rate (FDR) and present sharpened q-values for the FDR following Anderson (2008).[8] These are referenced in the tables as multiple-inference adjusted p-values.

[6]The table also shows that the training had little effect on literacy, which we pre-defined as a measurable outcome of training success. This is most likely because literacy was only covered in the initial three months of foundational skills training, and even there made up only one component of foundational skills, which also included numeracy and basic science.

[7]The pre-analysis plan can be accessed at https://www.socialscienceregistry.org/trials/4637.

[8]This approach was pre-specified in the pre-analysis plan, which divided outcomes into primary and secondary. We anticipated that the secondary outcomes would be part of the exploratory analysis and pre-specified that we would not adjust for multiple hypothesis testing across them, in order to preserve the statistical power of the primary hypothesis tests.

Our analysis of the gender norms index deviates from the pre-analysis plan in two ways. First, the pre-analysis plan specified that we would only estimate the effect of training on the gender norms of female study participants. Later, we decided to expand this analysis to include male participants as well, to test the hypothesis that observing their female peers' success in the training and labor market could affect their gender norms.[9] Our decision to expand the sample for this analysis to male participants necessitated a second deviation: we restricted the gender norms index to only include survey items that reflect norms about women's role in society, even when they are reported by a male respondent. Specifically, the originally specified version included a section on the respondent's experience of domestic violence, their desired age of marriage and fertility, and a measure of their confidence in their own cognitive ability.[10] These survey items can be interpreted as a measure of gender norms when reported by female respondents, but not necessarily when reported by male respondents. For example, a female respondent's desired age of fertility partly reflects her norms about when women should have children and how much control they should have over their reproduction. However, a male respondent's desired age of fertility is less relevant to our research question because men's reproductive behavior is less constrained by cultural norms, and because we are specifically interested in norms about women's role in society. We therefore exclude these items when calculating the gender norms index. However, in the appendix we report the originally pre-specified analysis that uses all pre-specified measures of gender norms and restricts the analysis to the female sub-sample and we find similar results.

Appendix Table A4 reports tests for balance on observable characteristics across the treatment and control groups. The results of these tests show that the groups are balanced on a wide range of characteristics, both from the pre-randomization intake interview and the baseline survey conducted by the implementing NGO.

[9] Note that this deviation is only relevant to the gender norms index – we pre-specified all other analyses to be conducted on the full sample of study participants.

[10] We measured this confidence by eliciting the respondent's belief about how they would perform on a "Raven's matrix" cognitive test.

Appendix Table A5 shows tests for selective attrition. Attrition was low overall at approximately 6 percent of the sample, reflecting a high response rate for a marginalized population with high rates of migration. However, attrition was slightly more frequent in the control group than the treatment group, most likely reflecting the treatment group's higher incentive to stay close to the training center, which may have outweighed other incentives to migrate (the estimated difference in attrition rates is approximately 3 percent). The imbalance in attrition rates across the treatment and control groups was particularly high in the two centers in Southern Kaduna (Kagoro and Mando), shown in Appendix Table A6. This region experienced a flare-up of violent conflict between herders and farmers that began in 2016 and lasted well into the second half of 2017.[11] This conflict led to substantial internal displacement in the region and likely increased people's incentive for outward migration (Ibrahim et al., 2022). The results in Appendix Table A7 show that in the other four centers not located in Southern Kaduna, attrition is balanced across the treatment and control group. A robustness test, reported in Appendix Table A8, shows that excluding the two centers in Southern Kaduna does not substantially change the estimated effects of training.

Appendix Table A4 also shows that the remaining sample of non-attriters is balanced on the relevant characteristics, so the imbalance created by the attrition process is unlikely to have a large effect on our estimates. Appendix Table A5 shows that there is imbalance in some characteristics between attriters and non-attriters, for example gender and Almajiri status. However, due to the low overall attrition rate, the remaining sample of non-attriters is still strongly balanced on all variables. In Appendix Table A9, we also verify that our estimates do not change substantially when we control for an extended set of observed characteristics that include the variables that were unbalanced across attriters from the treatment and control groups. To quantify the possible extent of bias from selective attrition, Appendix Table A10 reports Lee bounds, which do not include zero for any of our main outcomes of interest.

[11] https://www.chathamhouse.org/2017/02/violence-southern-kaduna-threatens-undermine-nigerias-democratic-stability.

# 6. Results

In this section, we present our results. We first discuss how the training affected economic outcomes, gender norms and social networks. We then present exploratory evidence on factors that contributed to the change i gender norms, and finally, show robustness checks.

## 6.1. Effects on economic outcomes

Table 1 examines our main economic outcomes. All outcomes in this table are mean effects indices calculated following Kling et al. (2007). The indices in the top panel were specified as primary outcomes in the pre-analysis plan, and for these outcomes, we present multiple inference adjusted p-values to account for multiple hypothesis testing (see discussion in Section 5).[12] The estimates in the top row shows that the training had substantial effects on income and employment, boosting this index by .24 standard deviations.

To delve further into these results, Appendix Table A11 presents effects on the individual components comprising this index. These estimates indicate large effects compared to the control mean. For example, the training increased wage employment by 39 percent and increased wage income by 54 percent.

The bottom panel of Table 1 also shows that the training led to a large positive effect on an index of assets and expenditures, which increased by 0.18 standard deviations. Appendix Table A12 presents estimates on the components of this index. We observe substantial increases in expenditures in almost all measured categories, including food, cell phone minutes, clothes, medical expenses, and leisure. In addition, we observe increased ownership of cell phones, generator sets, and bank accounts. The bottom row of Table 1 also shows a modest increase in job search behavior. Overall, these results indicate that the COSDEC

[12] In this table and all remaining tables, we adjust p-values for multiple inference across all primary outcomes, not just those shown in this table. The pre-analysis plan specified seven primary outcomes: income-generating activities, gender norms, caregiver gender norms, attitudes toward religious enforcement, participation in political and religious violence, anti-social behavior, and generosity toward other religions.

training had a transformative impact on the economic lives of participants, substantially expanding their economic opportunities.

In Appendix Table A13, we additionally examine how the training affected self-esteem and subjective well-being. The self-esteem index is composed of seven questions that measure the respondents feelings of self-worth and control over their life; while the well-being index is composed of a five-question mental health screening tool (MHI-5) and a measure of life-satisfaction (Cantril's ladder).[13] Results in this table indicate that the training did not affect self-esteem, but led to an increase in well-being. This effect is driven by a large increase in the Cantril's ladder item, which suggests that trained participants felt more satisfied with their life as a whole.

## 6.2. Effects on youth gender norms

Table 1 also shows that the training had a substantial effect on gender norms, increasing the youth gender norms index by 0.12 standard deviations and the caregiver gender norms index by 0.16 standard deviations. We unpack these impacts through two additional steps. In Table 2 we explore which aspects of youth gender norms shift in response to the training by grouping together questions on household decision-making, women's role in society and domestic violence, and presenting estimates on these component indices. In Appendix Table A1, we further present effects on all underlying individual questions that comprise these component indices.

Table 2 shows that the training had important effects on youths' norms about household decision-making and women's role in society, increasing the corresponding component indices by 0.12 and 0.10 standard deviations. And, Appendix Table A1 shows that the impacts on household decision making are driven by changes in views around whether women should have a say in their decisions to have children, and the household's lending and bor-

[13] Cantril's ladder asks: "Please imagine a ladder with steps numbered from zero at the bottom to 10 at the top. The top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible life for you. Which step are you currently on?"

rowing. For example, only 49 percent of individuals in the control group believe that women should have a say in the decision to have children; and exposure to the training increased this measure by 4.3 percentage points (a 9 percent increase above the control mean). Additionally, only 24 percent of individuals in the control group agreed that women should have a say in the household's decision to borrow money; and exposure to the training increased this measure by 5.7 percentage points (which represents a 24 percent increase). The training also reshaped respondents' norms about whether boys should do as much domestic work as girls, whether girls are as capable as boys in school, and whether girls should speak as much as boys in the classroom. For instance, only 28 percent of individuals in the control group believe that boys should do as much domestic work as girls, and the training increased this measure by almost 9 percentage points, or approximately 31 percent.

In contrast, the training had little effect on norms around domestic violence against women. Table 2 demonstrates that effects on this component index are small and statistically insignificant (as are the effects on the underlying variables in Appendix Table A1). This pattern of results suggests that exposure to the training, which is delivered in a school setting and provides participants with labor market opportunities, changes norms about related contexts such as education and financial decisions more than it changes norms about other contexts such as intra-household violence.

Appendix Table A14 shows results for the originally pre-specified index of gender norms of female participants, which includes the additional gender norms items that apply only to women (see discussion in Section 5). These results are very similar to those from our preferred specification in Table 2, suggesting a robust effect on youth gender norms as a result of the training.

While it is in principle possible that these estimated effects on gender norms suffer from social desirability bias, there are two reasons why we believe that this is not a major concern. First, the enumerators were locally recruited and therefore come from the same cultural

background as the respondents, where conservative gender norms are widespread. It is therefore unlikely that respondents feel the need to hide conservative gender norms. This is reflected in the large control means of our survey data, as discussed when interpreting the effect sizes above, and in Section 2.1. A few additional statistics from Appendix Table A1 lend further support to this idea: in the control group, 88 percent of respondents said that men should be the sole decision-maker about large household purchases; 77 percent said that the husband alone should decide whether his wife is allowed to work outside the home; and 49 percent said that it is more important to educate boys than girls. Second, the training had no explicit gender component, so gender issues would not have been any more salient to individuals who received training than to those who did not. It is therefore unlikely that social desirability would have led respondents who received the training to under-report conservative gender norms relative to respondents who did not receive the training.

Table 3 explores heterogeneous impacts by gender of the participants. For female participants, the training led to a statistically significant increase in the gender norms index of approximately 0.14 standard deviations. The effect on male participants is slightly smaller at 0.08 standard deviations, and not statistically significant. However, additional analysis shows that the difference in estimated effects between male and female participants is not statistically significant.[14] Thus, we cannot reject the hypothesis that the training affected male and female respondents equally, albeit this could partly be because we lack the statistical power to detect small differences in effects across the two sub-samples.

## 6.3. Effects on caregiver gender norms

We next turn toward examining if the training had effects on the norms of participants' caregivers. Caregiver norms may change for several different reasons. For instance, caregivers' views on women's work and education may shift upon observing their wards' positive experiences with the training, and their improved labor market outcomes. We expect this effect

[14] A Chow test does not reject the null-hypothesis of equal effects for male and female participants with a p-value of 0.12.

to emerge in the case of caregivers with female wards, specifically. Alternatively, caregivers may notice that religious authority figures such as Mallams did not protest to the program or shut it down, though it led to women participating in training and employment. From this they might infer that religious leaders are less opposed to women participating in these types of activities than previously assumed, which could lead to a change in their views through a correction of misperceived norms (Bursztyn et al., 2020). We expect this type of response to emerge for caregivers exposed to the program, regardless of whether their ward is male or female.

One difficulty in estimating this effect arises from the fact that some caregivers have multiple wards enrolled in the training, so they cannot be assigned to a unique stratum fixed effect (e.g. if different wards applied to different trades). In our baseline specification, we address this by restricting the sample to caregivers with a single ward in the study. For this subsample, we can estimate the effect of training with equation 1, the same equation we used to estimate the effect for the participating youth. Results of this estimation are presented in Table 4.

These results show that, as with the youth, the training had a substantial effect on caregiver's norms about the role of women in society, increasing the corresponding component index by 0.16 standard deviations. In addition, the training had a large effect on caregivers' norms about women's work decisions, increasing this component index by 0.12 standard deviations. We report the effect of training on all variables that make up the index in Appendix Table A2.

We cannot fully disentangle the mechanism through which this effect arises. However, we present some suggestive evidence in Appendix Table A15, which disaggregates the caregiver effect based on the gender of their ward. We observe that the effect is strongest for caregivers with female wards in the study. In contrast, the effect on caregivers with only male wards in the study is both small and statistically insignificant. This is consistent with a mechanism in

which observing a female ward's success in the training and subsequently the labor market leads to a change in the caregivers' gender views.

In Appendix Table A16, we additionally present results including caregivers who have multiple wards enrolled in the training. In this specification, we allow the caregivers to have multiple strata fixed effects – one for every strata in which they have a participating ward. These estimates are very similar to the ones from our baseline approach in Table 4.

The impact on caregiver norms is particularly important because of the considerable influence caregivers have on the work decisions of their female wards. As discussed above in Section 2, young women in Northern Nigeria generally have little autonomy over their work decisions, usually needing the permission of their husband, parents or other caregivers to seek employment. Many observers attribute the very low female labor force participation in the region to religious norms these caregivers hold against women working outside of the household (Egwurube, 2016; Enfield, 2019). By softening these norms, the COSDEC training could have substantial positive impacts on young women's ability to join the labor force in the future. As shown in Appendix Table A2, changes in the caregiver's gender norms reflect changes in views around whether women should occupy leadership positions in society and, as in the case of youth gender norms, whether boys should do as much domestic work as girls and whether girls should speak as much as boys in the classroom. These effects, combined with the observed impact on questions related to women's work decisions, reinforces the idea that the training alters gender norms pertaining to educational and economic opportunities.

## 6.4. Effects on social networks

In Table 5 we examine the training's effect on participants' social networks. We observe that participants experienced significant increases in the number of friends who are employed, friends of the opposite gender, and friends who live outside the neighborhood. It is likely that this broadening of social interactions is at least partly explained by the increase in labor

market opportunities of trained individuals (a hypothesis we explore in the next subsection below).

Appendix Table A17 explores whether the integration into new, broader social networks described above led to a loss of connections with the participants' old social networks. We find no evidence for this. The training has no effect on trust in the respondent's family, people in the respondent's own neighborhood or people the respondent knows personally. It also has no negative effect on the time participants spent with family or their probability of borrowing money from friends and family. Furthermore, it has no effect on the probability that the respondent lives or studies with a Mallam, which suggests that program participants were not excluded from their old religious networks. On the contrary, the training had positive effects on the respondent's confidence in religious leaders of their own religion and on the number of friends in religious school. In addition, individuals who received training were more likely to participate in youth groups and sports clubs. Overall, these results suggests that the broadening of social networks described above did not come at the cost of weakening existing social ties.

While social networks appear to have broadened along dimensions related to the labor market, we see more mixed evidence on network expansion along dimensions of religiosity. As can be seen in Appendix Table A18, people do not report more friendships from those of other religions. They do report greater trust of people and leaders from other religions, but there is also no effect on decisions to donate to charities from another religion (the Red Cross for Muslims and the Red Crescent for Christians).

In addition, in this table, we observe a decrease in an index of religiosity. However, the largest and most precise effects in the index stem from Mafita participants spending less time on religious activities and an increase in the belief that other things are more important than religion, respectively. This is consistent with Mafita participants working more and potentially attaching relatively greater value to work. More generally, this result points to

the opportunity cost of time playing a role in religious engagement. Appendix Table A19 also shows that the training had no additional downstream effects on religious enforcement or anti-social behavior.[15] Relatedly, we also do not observe systematic or consistent effects on participation in incidents of religious and political violence, or changes in attitudes toward political violence (see Table A20). Almost all coefficients are small and statistically insignificant with the exception of one, and display varied signs.

Overall, these results suggest that the training caused a substantial expansion in participants' social networks along dimensions related to employment, gender and geographic reach, without significantly reducing their existing social ties, or altering their views in other types of domains such as those related to political and religious conflict.

## 6.5. Exploring factors that contributed to the gender norms change

The results we have presented so far provide strong evidence that the Mafita vocational training program reshaped gender norms in a religiously conservative setting. We now turn toward exploring two key factors that may have contributed to the program's ability to engender these changes.

**The role of gender composition in the classroom.** The first potential factor is interaction between male and female participants in the classroom. As described in Section 3, the Mafita program trained male and female participants together in the same classroom, in a cultural context where interaction between unmarried people of the opposite gender is often restricted. It is possible that this interaction contributed to the training's impact on gender norms. For instance, boys in mixed-gender classrooms may have seen girls perform as well as boys, which may have led them to update their beliefs about gender roles. Conversely, girls may have noticed that they are as capable as boys and updated their gender norms in a similar way.

[15] As described in the pre-analysis plan, religious enforcement was partly measured through a series of vignettes that asked respondents whether they would take action to punish violations of religious norms like blasphemy or drinking alcohol.

If this kind of interaction played an important role in reshaping gender norms, we would expect the effect of the training to be larger in mixed-gender classrooms than in gender-segregated ones. We test whether this is the case by exploiting variation in the gender-composition of classrooms across training centers and trades.

As noted in Section 4.1, study participants applied to a specific center-trade combination and were randomized into either receiving the training there, or not receiving the training at all. At each center, there was one classroom for participants receiving training in a particular trade. And, in the course of the training, participants spent most of their time in this classroom with the other participants in the same trade and center.

The gender-mix of the trained participants varied substantially across the trade-center combinations, ranging from fully segregated to a nearly even split. We examine heterogeneity based on the gender-mix of the study participant's classroom. Specifically, we calculate a measure of the classroom's gender mix, defined as one minus the absolute value of the difference between the percentage of boys and girls in a classroom. For this calculation, we focus on the trade-center combination a study participant applied to, even if they were not chosen to be trained (i.e., were randomly assigned into the control group). This variable ranges from 0 for fully gender-segregated classrooms to 1 for classrooms with a perfectly even gender split. The sample mean is 0.33, meaning that the average classroom is split two-thirds to one-third between male and female participants, or vice versa. We then estimate regressions in which we interact the training indicator with this gender-mix variable.

The results of this analysis, presented in Appendix Table A21, show that the effect of training does not differ significantly between classrooms with a different gender-mix. For the gender norms outcome, the coefficient associated with the interaction term is small and not statistically significant, suggesting that the effect of training is not distinguishably different in gender-segregated and gender-mixed classrooms. We note that gender-mix is not randomly assigned, as study participants were free to apply to any trade and center, and thus

may be self-selecting into the gender composition of their (potential) training environment. This suggests that the interaction effects are not well-identified: they may reflect both the heterogeneous effect of being trained in classrooms with a different gender mix, and heterogeneity across the types of participants who applied to trades with a different gender mix. However, the most plausible forms of self-selection would likely lead us to find larger effects on gender norms based on the gender-mix of the classroom. For instance, it is possible that individuals who applied to less gender-mixed trades were more conservative and therefore less likely to change their gender norms.Thus, though this analysis should be taken as suggestive, it indicates that the gender mix of the classroom did not play a major role in determining the effect of training.

**The role of labor market success.** Another key feature of the Mafita program was the great care it took in designing a training curriculum adapted to local labor market conditions, and in ensuring a high quality of instruction (see Section 3). Perhaps because of this, the training led to large improvements in labor market outcomes that had a transformative impact on the lives of study participants, as shown in Section 6.1 above. It is possible that these changes in the economic lives of participants contributed to the effect of the training on gender norms, for example by demonstrating that women could succeed in the labor market, or broadening their social network. Before the training, study participants were part of relatively parochial social networks in poorer, more religiously conservative neighborhoods. Participants who found employment may have been exposed to a new set of beliefs through interactions with coworkers, customers and bosses, and these interactions may have led to a change in their gender norms.[16] This possibility is supported by our analysis of social networks in Section 6.4, which showed that training led to a substantial broadening of participants' social networks.

We now explore how much of the training's effect on gender norms and social networks

[16]The hypothesis that a widening of social interactions affected participant's gender norms is consistent with previous evidence that interactions with Muslims from other countries led to a change in gender norms among participants in the Haj (Clingingsmith et al., 2009).

could potentially arise from its success in creating labor market opportunities. To do this this, we conduct a mediation analysis that follows the approach of Heller et al. (2017). Table 6 presents these results. Column 1 shows the experimentally estimated effects of the training on the potential mediator: the income and employment index. This is the same estimate shown in Table 1, which indicates that the training increased income and employment by 0.24 standard deviations.

Column 2 shows the correlation between the income and employment index (the potential mediator) and the two outcomes of interest, gender norms and social networks, in the control group. Both correlations are shown to be strong: individuals in the control group who score higher on the employment and income index also have broader social networks and more equitable gender norms. While this correlation is not experimentally estimated, it is consistent with the hypothesis that employment is associated with a broadening of social networks and a shift away from conservative gender norms.

Column 3 reports the share of the total effect of training on the respective outcome that could be explained by the training's effect on income and employment. This estimate comes from multiplying the effect of the training on the income and employment index reported in column 1 by the link between the index and the respective outcome in column 2, and then dividing by the effect of the training on the outcome. Following Heller et al. (2017), we obtain confidence intervals by bootstrapping with 5000 bootstrap replications.

The results in column 3 show that the effect of the training on income and employment may explain close to 20 percent of its effect on both gender norms and social networks. While these results utilize non-experimental variation, and should therefore be interpreted somewhat cautiously, they do provide suggestive evidence that the increase in trained individuals' labor market opportunities played a meaningful role in expanding their social ties and altering their gender norms.

## 6.6. Robustness tests

In this section we conduct several additional robustness tests. First, we consider and present evidence against the concern that our results are driven by the effect of an add-on entrepreneurship training. As discussed in Section 6, this additional training was offered non-randomly to 352 individuals in the treatment group (approximately 41 percent of the group), as well as 19 individuals in the control group. We therefore cannot estimate its effect separately from the effect of the vocational training component of the Mafita program. However, we report regressions that exclude participants in this additional training in Appendix Table A22. These estimates show that the precision and magnitude of all our key outcomes remain unchanged, which indicates that the effects of the training on employment and gender norms are not driven by the additional entrepreneurship training and instead reflect the program's core vocational training component.

Next, we check the robustness of our results to alternate control sets in Appendix Table A9. In this table, the first column includes no controls; the second column is our baseline specification shown throughout the tables, which includes stratum fixed effects and basic demographic controls for participants (except when we measure gender norms of caregivers); and the third column includes additional control variables from the baseline survey. As discussed in Section 4.2, the baseline data is only available for 85% of the participants. We impute values of the baseline control variables using the group mean. The estimates from the first and third columns are very similar to the estimates from the baseline specification in the second column, which demonstrates that our results are insensitive to the control set.

Finally, we consider the sensitivity of our results to the way in which the indices are constructed. In Appendix Table A23 we present results for the economic and gender norms outcomes using inverse-covariance weighted versions of the outcome indices, following Anderson (2008). These estimates are also very similar to the baseline estimates in Table 1, which use unweighted indices following Kling et al. (2007). These checks help to verify that

the training exerts robust social and economic effects on the lives of program participants.

# 7. Conclusion

This paper provides evidence of the transformative potential of economic interventions in re-shaping gender norms, in the context of a religiously conservative society. Our evidence comes from a randomized controlled trial of the Mafita program in Northern Nigeria, which extended vocational training to youth from economically disadvantaged and religious backgrounds, including boys residing in Quranic boarding schools and girls receiving an Islamic education. Our findings reveal several significant insights.

First, vocational training led to substantial improvements in the economic well-being of participants, who had higher incomes and higher rates of employment a year after the training. This shows that skills training can significantly improve livelihoods even among highly disadvantaged populations with low attachment to the labor market.

Second, though the curriculum did not discuss female empowerment or address gender inequity in any way, it had a pronounced impact on gender norms. Participants who received training exhibited more progressive gender attitudes, supporting women's participation in household decisions, endorsing shared domestic responsibilities between boys and girls, and acknowledging the capabilities of girls in education.

Furthermore, the impact of the training extended beyond the participants themselves. Caregivers and guardians of female participants also exhibited improved gender attitudes, particularly with regard to supporting female labor force participation and women's roles in leadership and decision-making. Given the influential role of caregivers in Northern Nigerian society, this could be highly consequential for furthering gender equality in the region. One implication of this finding is that avoiding an explicit focus on gender empowerment or targeting women specifically may help to successfully re-shape gender attitudes in culturally conservative societies.

Our mediation analysis also suggests that increased employment served as a key pathway through which the training influenced gender norms. This underscores the interplay between economic development and gender norms, and the role economic improvements can play in spurring positive changes in societal norms.

In conclusion, our evidence highlights the potential of skills training as a powerful tool for promoting gender equity and economic empowerment, offering a path towards a more inclusive and equitable society in regions where traditional norms have persisted for generations. Our findings underscore the importance of continuing efforts to expand access to labor market skills, not only for their economic benefits but also for their potential to catalyze social change and challenge deeply ingrained gender inequalities.

# References

**Abou Daher, Chaza, Erica M Field, Kendal M Swanson, and Kate H Vyborny**, "Drivers of change: How intra-household preferences shape employment responses to gender reform," *NBER Working Paper*, 2023.

**Adeyemi, Oluwagbemiga E, Kolawole E Odusina, and Akinwole E Akintoye**, "Religion and labour force participation in Nigeria: is there any inequality among women?," *African journal of reproductive health*, 2016, *20* (3), 75–84.

**Adoho, Franck, Shubha Chakravarty, Dala T Korkoyah, Mattias KA Lundberg, and Afia Tasneem**, "The impact of an adolescent girls employment program: The EPAG project in Liberia," *World Bank Policy Research Working Paper*, 2014, (6832).

**Alesina, Alberto, Paola Giuliano, and Nathan Nunn**, "On the origins of gender roles: Women and the plough," *The quarterly journal of economics*, 2013, *128* (2), 469–530.

**Alfonsi, Livia, Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali**, "Tackling youth unemployment: Evidence from a labor market experiment in Uganda," *Econometrica*, 2020, *88* (6), 2369–2414.

**Almås, Ingvild, Alex Armand, Orazio Attanasio, and Pedro Carneiro**, "Measuring and changing control: Women's empowerment and targeted transfers," *The Economic Journal*, 2018, *128* (612), F609–F639.

**Anderson, Michael L**, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 2008, *103* (484), 1481–1495.

**Armand, Alex, Orazio Attanasio, Pedro Carneiro, and Valérie Lechene**, "The effect of gender-targeted conditional cash transfers on household expenditures: Evidence from a randomized experiment," *The Economic Journal*, 2020, *130* (631), 1875–1897.

**Ashraf, Nava, Natalie Bau, Corinne Low, and Kathleen McGinn**, "Negotiating a better future: How interpersonal skills facilitate intergenerational investment," *The Quarterly Journal of Economics*, 2020, *135* (2), 1095–1151.

**Attanasio, Orazio, Adriana Kugler, and Costas Meghir**, "Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial," *American Economic Journal: Applied Economics*, 2011, *3* (3), 188–220.

**_ , Arlen Guarín, Carlos Medina, and Costas Meghir**, "Vocational training for disadvantaged youth in Colombia: A long-term follow-up," *American Economic Journal: Applied Economics*, 2017, *9* (2), 131–143.

**Attanasio, Orazio P and Valérie Lechene**, "Efficient responses to targeted cash transfers," *Journal of political Economy*, 2014, *122* (1), 178–222.

**Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman**, "Women's empowerment in action: evidence from a randomized control trial in Africa," *American Economic Journal: Applied Economics*, 2020, *12* (1), 210–259.

**_ , Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman**, "Labor markets and poverty in village economies," *The Quarterly Journal of Economics*, 2017, *132* (2), 811–870.

**Bedoya, Guadalupe, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, and Jeremy P Shapiro**, "No household left behind: Afghanistan targeting the ultra poor impact evaluation," *NBER Working Paper*, 2019.

**Bertrand, Marianne, Emir Kamenica, and Jessica Pan**, "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 2015, *130* (2), 571–614.

**Blattman, Christopher and Jeannie Annan**, "Can employment reduce lawlessness and rebellion? A field experiment with high-risk men in a fragile state," *American Political*

*Science Review*, 2016, *110* (1), 1–17.

_ **and Laura Ralston**, "Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs," *Available at SSRN 2622220*, 2015.

**Blumenstock, Joshua, Oeindrila Dube, and Karrar Hussain**, "Can Secular Media Create Religious Backlash? Evidence from Pakistan's Media Liberalization," 2022.

**Bobonis, Gustavo J**, "Is the allocation of resources within the household efficient? New evidence from a randomized experiment," *Journal of political Economy*, 2009, *117* (3), 453–503.

**Bossuroy, Thomas, Markus Goldstein, Bassirou Karimou, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Catherine C Thomas, Christopher Udry, Julia Vaillant et al.**, "Tackling psychosocial and capital constraints to alleviate poverty," *Nature*, 2022, *605* (7909), 291–297.

**Buchmann, Nina, Erica Field, Rachel Glennerster, Shahana Nazneen, and Xiao Yu Wang**, "A signal to end child marriage: Theory and experimental evidence from Bangladesh," *American Economic Review*, 2023, *113* (10), 2645–2688.

**Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott**, "Misperceived social norms: Women working outside the home in Saudi Arabia," *American Economic Review*, 2020, *110* (10), 2997–3029.

_ **, Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott**, "How Are Gender Norms Perceived?," *NBER Working Paper*, 2023.

**Callaway, Barbara J**, "Law Education and Social Change: Implications for Hausa Muslim Women in Nigeria," in Lynne B. Iglitzin and Ruth Ross, eds., *Women in the World, 1975–1985: The Women's Decade*, Santa Barbara: ABD-Clio, 1986.

**Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas-Shady, and Yuri Soares**, "The labor market impacts of youth training in the Dominican Republic," *Journal of Labor Economics*, 2011, *29* (2), 267–300.

**Carranza, Eliana and David McKenzie**, "Job training and job search assistance policies in developing countries," *Journal of Economic Perspectives*, 2024, *38* (1), 221–244.

**Cho, Yoonyoung, Davie Kalomba, Ahmed Mushfiq Mobarak, and Victor Orozco-Olvera**, "Gender differences in the effects of vocational training: Constraints on women and drop-out behavior," *World Bank Policy Research Working Paper*, 2013, (6545).

**Clingingsmith, David, Asim Ijaz Khwaja, and Michael Kremer**, "Estimating the impact of the Hajj: religion and tolerance in Islam's global gathering," *The Quarterly Journal of Economics*, 2009, *124* (3), 1133–1170.

**Csapo, Marg**, "Religious, social and economic factors hindering the education of girls in northern Nigeria," *Comparative education*, 1981, *17* (3), 311–319.

**Deininger, Klaus, Aparajita Goyal, and Hari Nagarajan**, "Women's inheritance rights and intergenerational transmission of resources in India," *Journal of Human Resources*, 2013, *48* (1), 114–141.

**Dhar, Diva, Tarun Jain, and Seema Jayachandran**, "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India," *American economic review*, 2022, *112* (3), 899–927.

**Duflo, Esther**, "Women empowerment and economic development," *Journal of Economic Literature*, 2012, *50* (4), 1051–1079.

**Edmonds, Eric, Ben Feigenberg, and Jessica Leight**, "Advancing the agency of adolescent girls," *Review of Economics and Statistics*, 2023, *105* (4), 852–866.

**Egbue, Ngozi G**, "Gender division of domestic roles: Implications for social equality in Nigeria," *The Nigerian Journal of Sociology and Anthropology Vol*, 2012, *8*, 14.

**Egwurube, Joseph**, "Challenges facing women empowerment in contemporary Nigeria," *Miroirs: Revue des civilisations anglophone, ibérique et ibéro-américaine*, 2016, *2* (4), 134–168.

**Enfield, Sue**, "Gender roles and inequalities in the Nigerian labour market," Technical Report, IDS 2019.

**Erulkar, Annabel and Mairo Bello**, "The experience of married adolescent girls in northern Nigeria," Technical Report, Population Council 2007.

**Fernández, Raquel**, "Cultural change as learning: The evolution of female labor force participation over a century," *American Economic Review*, 2013, *103* (1), 472–500.

**_ and Alessandra Fogli**, "Culture: An empirical investigation of beliefs, work, and fertility," *American economic journal: Macroeconomics*, 2009, *1* (1), 146–177.

**Ferrara, Eliana La, Alberto Chong, and Suzanne Duryea**, "Soap operas and fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 2012, *4* (4), 1–31.

**Field, Erica and Kate Vyborny**, "Women's mobility and labor supply: experimental evidence from Pakistan," *Asian Development Bank Economics Working Paper Series*, 2022, (655).

**Field, Erica M, Leigh L Linden, Ofer Malamud, Daniel Rubenson, and Shing-Yi Wang**, "Does vocational education work? Evidence from a randomized experiment in Mongolia," *NBER Working Paper*, 2019.

**Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore**, "On her own account: How strengthening women's financial control impacts labor supply and gender norms," *American Economic Review*, 2021, *111* (7), 2342–2375.

**Gazeaud, Jules, Nausheen Khan, Eric Mvukiyehe, and Olivier Sterck**, "With or without him? Experimental evidence on cash grants and gender-sensitive trainings in Tunisia," *Journal of Development Economics*, 2023, *165*, 103169.

**Glennerster, Rachel, Joanna Murray, and Victor Pouliquen**, "The media or the message? Experimental evidence on mass media and modern contraception uptake in Burkina Faso," *Centre for the Study of African Economies Working Paper*, 2021.

**Haerpfer, C., R. Inglehart, A. Moreno, C. Welzel, K. Kizilova, Diez-Medrano J., M. Lagos, P. Norris, E. Ponarin, and B. Puranen**, "World Values Survey: Round Seven – Country-Pooled Datafile," Technical Report, JD Systems Institute I& WVSA Secretariat, Madrid, Spain I& Vienna, Austria 2020.

**Hasan, Amer**, "Time allocation in rural households: The indirect effects of conditional cash transfer programs," *World Bank Policy Research Working Paper*, 2010, (5256).

**Heller, Sara B, Anuj K Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A Pollack**, "Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago," *The Quarterly Journal of Economics*, 2017, *132* (1), 1–54.

**Hirshleifer, Sarojini, David McKenzie, Rita Almeida, and Cristobal Ridao-Cano**, "The impact of vocational training for the unemployed: experimental evidence from Turkey," *The Economic Journal*, 2016, *126* (597), 2115–2146.

**Ibrahim, Aliyu H, Nura Alhaji Yaro, Murjanatu Yahaya Ibrahim, and Yusuf Adama**, "Evaluating the causes and consequences of internal displacement of persons in Kaduna State," *FUDMA journal of sciences*, 2022, *6* (2), 120–126.

**ICF**, "Demographic and Health Surveys (various)," Technical Report, Funded by USAID, Rockville, Maryland: ICF 2004-2017.

**Jayachandran, Seema**, "The roots of gender inequality in developing countries," *Annual Review of Economics*, 2015, *7* (1), 63–88.

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

**Jensen, Robert**, "Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India," *The Quarterly Journal of Economics*, 2012, *127* (2), 753–792.

_ **and Emily Oster**, "The power of TV: Cable television and women's status in India," *The Quarterly Journal of Economics*, 2009, *124* (3), 1057–1094.

**Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz**, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119.

**Laouan, Fatouma Zara**, "CARE Rapid Gender and GBV Assessment," Technical Report, CARE Nigeria 2018.

**McKelway, Madeline**, "Women's self-efficacy and women's employment: Experimental evidence from India," *Unpublished Working Paper*, 2018.

**Newton, Danielle, Andrew Day, Margaret Giles, Joanne Wodak, Joe Graffam, and Eileen Baldry**, "The impact of vocational education and training programs on recidivism: A systematic review of current experimental evidence," *International journal of offender therapy and comparative criminology*, 2018, *62* (1), 187–207.

**Okojie, Christiana**, "Achieving selfreliance in food production in Nigeria: Maximising the contribution of rural Women," *Journal of social development in Africa*, 1991, *6* (2), 33–52.

**Onwutuebe, CJ**, "Religious interpretations, gender discrimination and politics in Africa: A case study of Nigeria," 2013.

**Open Data for Africa**, "Nigeria Data Portal," Technical Report, Open Data for Africa 2022. Accessed: 2022-09-27.

**Ostien, Philip**, "The Muslim majority in northern Nigeria," *Creed & grievance: Muslim-Christian relations & conflict resolution in Northern Nigeria*, 2018, *11*, 37.

**Para-Mallam, Funmi J**, "Promoting gender equality in the context of Nigerian cultural and religious expression: beyond increasing female access to education," *Compare*, 2010, *40* (4), 459–477.

**Robson, Elsbeth**, "Wife seclusion and the spatial praxis of gender ideology in Nigerian Hausaland," *Gender, Place and Culture: A Journal of Feminist Geography*, 2000, *7* (2), 179–199.

**Seff, Ilana, Shadrack Steven, Alli Gillespie, Hannah Brumbaum, Hannah Kluender, Cassondra Puls, Andrea Koris, Veronica Akika, Julianne Deitch, and Lindsay Stark**, "A Family-Focused, Sibling-Synchronous Intervention in Borno State, Nigeria: Exploring the Impact on Family Functioning and Household Gender Roles," *Journal of Family Violence*, 2022, pp. 1–13.

**Zakaria, Yakubu**, "Entrepreneurs at home: Secluded Muslim women and hidden economic activities in Northern Nigeria," *Nordic Journal of African Studies*, 2001, *10* (1), 17–17.

Table 1: Impact on economic outcomes and gender norms

| | Mafita | N |
|---|---|---|
| *Primary* | | |
| Income and employment | 0.241*** | 1706 |
| | (0.046) | |
| | [0.001] | |
| Gender norms index (youth) | 0.120*** | 1706 |
| | (0.043) | |
| | [0.018] | |
| Gender norms index (caregivers) | 0.163** | 792 |
| | (0.073) | |
| | [0.048] | |
| *Secondary* | | |
| Assets and expenditures | 0.177*** | 1706 |
| | (0.047) | |
| Job search behaviour | 0.084* | 1706 |
| | (0.048) | |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). All outcome variables are mean effect indices, standardized using the control group mean and standard deviation. Each row presents results from a separate regression. All regressions control for individual characteristics at the time of randomization (age, gender, ethnicity, religion, and type of vulnerable group) and strata fixed effects, except for the caregiver gender norms index, where we only control for strata fixed effects. Robust standard errors are in parentheses. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 2: Examining effects on the gender norms of youth

| | Mafita |
|---|---|
| Gender norms index (youth) | 0.120*** |
| | (0.043) |
| *Components of index* | |
| Household decision-making | 0.120*** |
| | (0.043) |
| Role in society | 0.101** |
| | (0.043) |
| Domestic violence | 0.011 |
| | (0.047) |
| Number of youth | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). Each row presents results from a separate regression. The variable in the top panel is a mean effect index comprised of all gender norms outcomes. The bottom panel presents component mean effect indices that aggregate gender norms outcomes related to household decision-making, women's role in society, and domestic violence separately. All regressions control for individual characteristics at the time of randomization (age, gender, ethnicity, religion, and type of vulnerable group) and strata fixed effects. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 3: Heterogeneous effects by gender of participant

| | Mafita |
|---|---|
| *Male* | |
| Income and employment | 0.236*** |
| | (0.067) |
| Gender norms index (youth) | 0.083 |
| | (0.065) |
| Number of youth | 858 |
| *Female* | |
| Income and employment | 0.251*** |
| | (0.067) |
| Gender norms index (youth) | 0.137** |
| | (0.064) |
| Number of youth | 843 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT), separately for males in the first panel and females in the second panel. Each row presents results from a separate regression. All outcome variables are mean effect indices, standardized using the control group mean and standard deviation. All regressions control for strata fixed effects and demographics at the time of randomization (age, gender, ethnicity, religion, and type of vulnerable group). Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 4: Impact on caregiver gender norms

| | Mafita |
|---|---|
| Gender norms index (caregivers) | 0.163** |
| | (0.073) |
| *Components of index* | |
| Female work decisions | 0.122* |
| | (0.072) |
| Marriage and fertility | -0.013 |
| | (0.064) |
| Role in society | 0.160** |
| | (0.071) |
| Household decision-making | 0.021 |
| | (0.073) |
| Number of caregivers | 792 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). Each row presents results from a separate regression. The variable in the top panel is a mean effect index comprised of all caregiver gender norms outcomes. The bottom panel presents component mean effect indices that aggregate caregiver gender norms outcomes related to female work decisions, marriage and fertility, women's role in society, and household decision-making separately. We control for strata fixed effects. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 5: Impact on social networks

| | Control Mean | Mafita |
|---|---|---|
| Social network index | - | 0.201*** |
| | - | (0.048) |
| *Components of index* | | |
| Employed friends | 2.816 | 0.273*** |
| | (1.623) | (0.076) |
| Friends outside neighbourhood | 1.214 | 0.117* |
| | (1.343) | (0.065) |
| Friends of opposite gender | 0.045 | 0.013* |
| | (0.155) | (0.007) |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). Each row presents results from a separate regression. The outcome variable in the top panel is a mean effect index, standardized using the control group mean and standard deviation. The variables in the bottom panel are the individual outcomes that constitute this index. We control for strata fixed effects and demographics at the time of randomization (age, gender, ethnicity, religion, and type of vulnerable group). Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table 6: The role of income-generating activities on expanding social networks and gender norms

| | Col 1 | Col 2 | Col 3 |
|---|---|---|---|
| Gender norms index (youth) | 0.241*** | 0.083** | 0.166** |
| | (0.047) | (0.038) | [0.016, 0.585] |
| Social network index | 0.241*** | 0.165*** | 0.197** |
| | (0.046) | (0.036) | [0.092, 0.397] |

*Notes:* Col 1 reports the treatment effect of the program on the mediator, and Col 2 reports the association between the outcome and the mediator for the control group only. Col 3 reports the product of Col 1 and Col 2, divided by the treatment effect of the program on the outcome (not shown here). Col 3 is the share of the treatment effect explained by the mediator. Robust standard errors are in parentheses. Bootstrapped percentile-based 95% confidence intervals are in brackets. We bootstrap over 5000 samples. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

# A. Online Appendix

Table A1: Impact on components of gender norms index

| | Control Mean | Control SD | Mafita |
|---|---|---|---|
| *Household decision-making* | | | |
| Making large household purchases | 0.12 | 0.32 | 0.016 |
| | | | (0.016) |
| Making daily household purchases | 0.27 | 0.44 | 0.027 |
| | | | (0.021) |
| Wife's personal purchases | 0.30 | 0.46 | 0.014 |
| | | | (0.022) |
| Borrowing money | 0.24 | 0.42 | 0.057*** |
| | | | (0.021) |
| Lending money | 0.27 | 0.45 | 0.058*** |
| | | | (0.022) |
| Wife's choice of occupation | 0.26 | 0.44 | 0.024 |
| | | | (0.021) |
| Wife's place of work (home vs outside) | 0.23 | 0.42 | 0.008 |
| | | | (0.020) |
| Wife's working hours | 0.31 | 0.46 | 0.031 |
| | | | (0.021) |
| Wife's participation in groups | 0.21 | 0.41 | -0.002 |
| | | | (0.019) |
| Having children | 0.49 | 0.50 | 0.043* |
| | | | (0.024) |
| Money within household | 0.07 | 0.25 | 0.010 |
| | | | (0.012) |
| *Role in society* | | | |
| Daughters work outside of home | 0.78 | 0.41 | 0.027 |
| | | | (0.019) |
| Unmarried woman with job to be pitied (disagree) | 0.23 | 0.42 | -0.036* |
| | | | (0.020) |
| Women occupy leadership positions | 0.67 | 0.47 | 0.022 |
| | | | (0.022) |
| Women allowed to work outside of home | 0.77 | 0.42 | 0.021 |
| | | | (0.019) |
| Educating boys more important than educating girls (disagree) | 0.52 | 0.50 | -0.011 |
| | | | (0.023) |
| Boys should do as much domestic work as girls | 0.28 | 0.45 | 0.088*** |
| | | | (0.021) |
| A girl must obey her brother, even if he is younger (disagree) | 0.32 | 0.47 | -0.032 |
| | | | (0.022) |
| Girls should speak as much as boys in classrooms | 0.45 | 0.50 | 0.062*** |
| | | | (0.023) |
| Girls as capable as boys in school | 0.77 | 0.42 | 0.063*** |
| | | | (0.018) |
| *Domestic violence* | | | |
| Beat wife if she goes out without telling him (disagree) | 0.83 | 0.37 | -0.001 |
| | | | (0.018) |
| Beat wife if she neglects the children (disagree) | 0.80 | 0.40 | 0.003 |
| | | | (0.019) |
| Beat wife if she argues with him (disagree) | 0.82 | 0.39 | -0.019 |
| | | | (0.019) |
| Beat wife if she refuses to sleep with him (disagree) | 0.80 | 0.40 | 0.003 |
| | | | (0.019) |
| Beat wife if she refuses to cook (disagree) | 0.83 | 0.37 | 0.008 |
| | | | (0.018) |
| Beat wife if she refuses to clean the house (disagree) | 0.83 | 0.37 | -0.023 |
| | | | (0.018) |
| Beat wife if he suspects her of being unfaithful (disagree) | 0.76 | 0.43 | -0.021 |
| | | | (0.021) |
| Wife has right to disagree with her husband | 0.66 | 0.48 | 0.019 |
| | | | (0.022) |
| Number of youth | | | 1706 |

*Notes:* This table reports estimates of the effect of training on all components of the gender norms index for youth. Control means represent proportion of the sample, with higher values indicating more favorable attitudes toward women's rights. In the third column (Mafita), each row presents results from a separate regression. All regressions control for individual characteristics at the time of randomization (age, gender, ethnicity, religion, and type of vulnerable group) and strata fixed effects. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A2: Impact on components of caregiver gender norms index

| | Control Mean | Control SD | Mafita |
|---|---|---|---|
| *Female work decisions* | | | |
| 14y female single working at home | 0.70 | 0.46 | 0.036 |
| | | | (0.027) |
| 14y female single working outside home | 0.37 | 0.48 | 0.043 |
| | | | (0.033) |
| 18y female single working at home | 0.95 | 0.23 | 0.019 |
| | | | (0.015) |
| 18y female single working outside home | 0.68 | 0.47 | 0.047 |
| | | | (0.032) |
| 22y female single working at home | 0.99 | 0.12 | 0.006 |
| | | | (0.008) |
| 22y female single working outside home | 0.82 | 0.38 | 0.026 |
| | | | (0.027) |
| 14y female married working at home | 0.97 | 0.18 | 0.006 |
| | | | (0.013) |
| 14y female married working outside home | 0.47 | 0.50 | 0.044 |
| | | | (0.036) |
| 18y female married working at home | 0.99 | 0.10 | 0.007 |
| | | | (0.007) |
| 18y female married working outside home | 0.72 | 0.45 | 0.001 |
| | | | (0.032) |
| *Marriage and fertility* | | | |
| Women suitable marriage age | -19.18 | 2.96 | -0.049 |
| | | | (0.185) |
| Woman suitable age first child | -20.68 | 2.78 | -0.024 |
| | | | (0.190) |
| *Role in society* | | | |
| Unmarried woman with job to be pitied (disagree) | 0.17 | 0.37 | 0.003 |
| | | | (0.025) |
| Women should occupy leadership positions in society | 0.85 | 0.35 | 0.065*** |
| | | | (0.022) |
| Women should be allowed to work outside of home | 0.90 | 0.30 | 0.025 |
| | | | (0.021) |
| Educating boys is more important than educating girls (disagree) | 0.82 | 0.38 | -0.032 |
| | | | (0.029) |
| Boys should do as much domestic work as girls | 0.53 | 0.50 | 0.066* |
| | | | (0.034) |
| A girl must obey her brother, even if he is younger (disagree) | 0.58 | 0.49 | 0.018 |
| | | | (0.035) |
| Girls should speak as much as boys in the classroom | 0.75 | 0.43 | 0.069** |
| | | | (0.029) |
| Girls capable of doing as well as boys in school | 0.99 | 0.07 | 0.004 |
| | | | (0.004) |
| *Household decision-making* | | | |
| Making large household purchases | 0.92 | 0.27 | 0.005 |
| | | | (0.019) |
| Making daily household purchases | 0.85 | 0.36 | -0.017 |
| | | | (0.026) |
| Wife's personal purchases | 0.66 | 0.48 | 0.003 |
| | | | (0.034) |
| Borrowing money | 0.67 | 0.47 | 0.006 |
| | | | (0.033) |
| Lending money | 0.66 | 0.48 | 0.020 |
| | | | (0.033) |
| Wife's choice of occupation | 0.55 | 0.50 | 0.031 |
| | | | (0.036) |
| Wife's place of work (home vs outside) | 0.66 | 0.48 | 0.023 |
| | | | (0.034) |
| Wife's working hours | 0.63 | 0.48 | 0.012 |
| | | | (0.034) |
| Wife's participation in groups | 0.82 | 0.38 | 0.027 |
| | | | (0.027) |
| Having children | 0.45 | 0.50 | -0.061* |
| | | | (0.035) |
| Number of caregivers | | | 792 |

*Notes:* The table reports estimates of the effect of training on all components of the gender norms index for caregivers. In the third column (Mafita), each row presents results from a separate regression. All regressions control for strata fixed effects. Apart from the 'marriage and fertility' sub-index, means represent proportion of the sample. Higher values indicate more favorable attitudes toward women's rights. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A3: Impact on completion of training

| | Training | Literacy |
|---|---|---|
| Treatment | 0.670*** | 0.014 |
| | (0.017) | (0.048) |
| Number of youth | 1706 | 1706 |

*Notes:* This table reports the estimated effects of assignment to treatment (ITT). We include strata fixed effects. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A4: Balance on covariates for respondents

| | Control | Treatment | p-value |
|---|---|---|---|
| *Randomization* | | | |
| Age | 19.0 | 18.9 | 0.42 |
| Almajiri (%) | 24.1 | 25.4 | 0.54 |
| IQE (%) | 22.3 | 24.7 | 0.25 |
| Orphans/vulnerable children (%) | 26.9 | 24.8 | 0.32 |
| Early school leavers (%) | 21.7 | 21.0 | 0.70 |
| Person with disabilities (%) | 5.0 | 4.2 | 0.45 |
| *Baseline* | | | |
| Female (%) | 44.2 | 42.4 | 0.46 |
| Age | 16.6 | 16.3 | 0.47 |
| Hausa (%) | 70.0 | 68.3 | 0.45 |
| Muslim (%) | 72.0 | 71.6 | 0.85 |
| Hours studying religious texts per day | 3.0 | 3.1 | 0.49 |
| Taught trade (%) | 14.0 | 15.4 | 0.44 |
| Attended skills training (%) | 11.2 | 10.1 | 0.47 |
| Currently working (%) | 10.5 | 9.5 | 0.51 |
| Looking for work (%) | 53.7 | 54.1 | 0.86 |
| Currently an apprentice (%) | 7.1 | 6.4 | 0.58 |
| Number of youth | 847 | 859 | 1706 |

*Notes:* This table reports sample means of demographics at the pre-randomization intake survey and at baseline for the control and treatment groups. The last column reports p-values from a test of equality of means. The sample includes participants who took the survey at endline.

Table A5: Balance on covariates for attriters

| | Control | Treatment | p-value |
|---|---|---|---|
| *Randomization* | | | |
| Age | 19.2 | 19.7 | 0.32 |
| Almajiri (%) | 18.9 | 39.0 | 0.02 |
| IQE (%) | 16.2 | 4.9 | 0.08 |
| Orphans/vulnerable children (%) | 28.4 | 22.0 | 0.46 |
| Early school leavers (%) | 31.1 | 26.8 | 0.64 |
| Person with disabilities (%) | 5.4 | 7.3 | 0.68 |
| *Baseline* | | | |
| Female (%) | 59.3 | 23.1 | 0.00 |
| Age | 18.9 | 20.3 | 0.02 |
| Hausa (%) | 65.4 | 84.0 | 0.09 |
| Muslim (%) | 63.0 | 80.8 | 0.11 |
| Hours studying religious texts per day | 3.5 | 4.6 | 0.14 |
| Taught trade (%) | 9.3 | 26.9 | 0.04 |
| Attended skills training (%) | 20.4 | 19.2 | 0.91 |
| Currently working (%) | 5.6 | 17.4 | 0.10 |
| Looking for work (%) | 71.7 | 68.0 | 0.74 |
| Currently an apprentice (%) | 6.1 | 4.5 | 0.79 |
| Number of youth | 77 | 41 | 118 |

*Notes:* This table reports sample means of demographics at the pre-randomization intake survey and at baseline. The last column reports p-values from a test of equality of means. The sample includes participants who did not take the survey at endline.

Table A6: Attrition rates

| | Control | Treatment | p-value |
|---|---|---|---|
| Attrition rate (%) | 8.0 | 4.6 | 0.00 |
| Attrition rate - southern Kaduna (%) | 8.9 | 3.1 | 0.00 |
| Attrition rate - other centers (%) | 7.4 | 5.6 | 0.21 |

*Notes:* This table reports attrition rates across treatment and control groups and p-values from a test of equality of means.

Table A7: Balance on covariates for attriters, excluding the southern Kaduna centers

| | Control | Treatment | p-value |
|---|---|---|---|
| *Randomization* | | | |
| Age | 18.6 | 19.4 | 0.18 |
| Almajiri (%) | 31.7 | 46.7 | 0.21 |
| IQE (%) | 14.6 | 6.7 | 0.30 |
| Orphans/vulnerable children (%) | 19.5 | 20.0 | 0.96 |
| Early school leavers (%) | 34.1 | 20.0 | 0.20 |
| Person with disabilities (%) | 0.0 | 6.7 | 0.10 |
| *Baseline* | | | |
| Female (%) | 60.0 | 11.1 | 0.00 |
| Age | 18.0 | 20.3 | 0.01 |
| Hausa (%) | 96.0 | 100.0 | 0.42 |
| Muslim (%) | 96.0 | 100.0 | 0.40 |
| Hours studying religious texts per day | 5.0 | 5.8 | 0.31 |
| Taught trade (%) | 4.0 | 33.3 | 0.01 |
| Attended skills training (%) | 20.0 | 22.2 | 0.86 |
| Currently working (%) | 12.0 | 25.0 | 0.29 |
| Looking for work (%) | 79.2 | 66.7 | 0.37 |
| Currently an apprentice (%) | 4.2 | 6.7 | 0.74 |
| Number of youth | 44 | 30 | 74 |

*Notes:* This table reports sample means of demographics at the pre-randomization intake survey and at baseline for the control and treatment groups in all training centers except the two in Southern Kaduna. The last column reports p-values from a test of equality of means. The sample includes participants who took the survey at endline.

Table A8: Impact on economic outcomes and gender norms, excluding the southern Kaduna centers

| | Mafita | N |
|---|---|---|
| *Primary* | | |
| Income and employment | 0.227*** | 1019 |
| | (0.056) | |
| | [0.001] | |
| Gender norms index (youth) | 0.122** | 1019 |
| | (0.055) | |
| | [0.093] | |
| Gender norms index (caregivers) | 0.156 | 409 |
| | (0.106) | |
| | [0.197] | |
| *Secondary* | | |
| Assets and expenditures | 0.123** | 1019 |
| | (0.062) | |
| Job search behaviour | 0.028 | 1019 |
| | (0.057) | |

*Notes:* This table reports estimated effects of assignment to treatment (ITT) on respondents in all training centers except the two in Southern Kaduna. All outcome variables are mean effect indices, standardized using the control group mean and standard deviation. For the youth sample, we control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. For the caregiver sample, we control for strata fixed effects. Robust standard errors are in parentheses. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table A9: Impact on economic outcomes and gender norms, with additional controls

| | Mafita | Mafita | Mafita |
|---|---|---|---|
| *Primary* | | | |
| Income and employment | 0.248*** | 0.241*** | 0.243*** |
| | (0.047) | (0.046) | (0.046) |
| Gender norms index (youth) | 0.100** | 0.120*** | 0.122*** |
| | (0.046) | (0.043) | (0.043) |
| Gender norms index (caregivers) | 0.163** | 0.163** | 0.163** |
| | (0.073) | (0.073) | (0.073) |
| *Secondary* | | | |
| Assets and expenditures | 0.176*** | 0.177*** | 0.178*** |
| | (0.047) | (0.047) | (0.047) |
| Job search behaviour | 0.082* | 0.084* | 0.085* |
| | (0.048) | (0.048) | (0.047) |
| Social network index | 0.189*** | 0.201*** | 0.201*** |
| | (0.049) | (0.048) | (0.048) |
| Controls | No controls | Randomization | Randomization + Baseline |
| Number of youth | 1706 | 1706 | 1706 |
| Number of caregivers | 792 | 792 | 792 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). All outcome variables are mean effect indices, standardized using the control group mean and standard deviation. For the youth sample, there are no controls in the first column, we control for demographics at randomization and strata fixed effects in the second column, and we add demographics at baseline in the third column. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. For the caregiver sample, we control for strata fixed effects. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A10: Lee bounds

| | Lower | Upper | N |
|---|---|---|---|
| Income and employment | 0.131 | 0.284 | 1706 |
| Religious enforcement | -0.076 | 0.065 | 1706 |
| Gender norms index (youth) | 0.031 | 0.191 | 1706 |
| Anti-social behaviour | -0.099 | 0.084 | 1706 |
| Political and religious violence | -0.101 | 0.052 | 1706 |
| Donation experiment | -0.003 | 0.343 | 1706 |
| Attitudes toward political violence | -0.128 | -0.003 | 1706 |
| Trust in other religions | 0.002 | 0.279 | 1706 |
| Religiosity | -0.149 | 0.015 | 1706 |
| Well-being | 0.044 | 0.215 | 1706 |
| Self-esteem | -0.096 | 0.052 | 1706 |
| Literacy | -0.362 | 0.078 | 1706 |
| Assets and expenditures | 0.059 | 0.228 | 1706 |
| Social network index | 0.083 | 0.244 | 1706 |
| Job search behaviour | -0.050 | 0.113 | 1706 |

*Notes:* This table reports lower and upper bounds of the estimated effects of assignment to treatment (ITT), after accounting for attrition. All outcome variables are mean effect indices, standardized using the control group mean and standard deviation.

Table A11: Impact on income and employment outcomes

| | Control Mean | Mafita |
|---|---|---|
| Wage-employed work in past month | 0.109 | 0.042*** |
| | (0.311) | (0.016) |
| Self-employed work in past month | 0.380 | 0.137*** |
| | (0.486) | (0.024) |
| Days of wage-employed work in past month | 1.547 | 0.359 |
| | (5.367) | (0.263) |
| Days of self-employed work in past month | 6.187 | 1.942*** |
| | (9.495) | (0.474) |
| Income from wage-employed work in past month | 752.893 | 403.365** |
| | (3043.890) | (174.867) |
| Profit from business activities in past month | 2573.583 | 997.109*** |
| | (6054.716) | (318.240) |
| Number of youth | | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). We control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. Standard deviations of the control group are in parentheses, below the control means. Robust standard errors are in parentheses, below the ITT estimates. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A12: Impact on assets and expenditures

| | Control Mean | Mafita |
|---|---|---|
| Do you own an electric iron? | 0.649 | 0.035 |
| | (0.477) | (0.022) |
| Do you own a fan? | 0.616 | 0.000 |
| | (0.486) | (0.022) |
| Do you own a television? | 0.603 | 0.025 |
| | (0.489) | (0.023) |
| Do you own a refrigerator? | 0.282 | 0.019 |
| | (0.450) | (0.021) |
| Do you own a generating set? | 0.157 | 0.047*** |
| | (0.364) | (0.018) |
| Do you own a satellite or cable TV? | 0.238 | 0.004 |
| | (0.426) | (0.021) |
| Do you have electricity? | 0.882 | 0.010 |
| | (0.323) | (0.015) |
| Do you have a bank account? | 0.758 | 0.037* |
| | (0.427) | (0.020) |
| Do you own a cell phone? | 0.982 | 0.014*** |
| | (0.132) | (0.005) |
| (Expenses) Food | 504.746 | 103.393** |
| | (959.312) | (50.779) |
| (Expenses) Transport | 296.139 | 60.482** |
| | (542.808) | (27.943) |
| (Expenses) Airtime and data for phone | 169.818 | 39.410** |
| | (297.854) | (16.449) |
| (Expenses) Medical | 535.331 | 241.560*** |
| | (1942.483) | (93.077) |
| (Expenses) Clothes | 1017.766 | 288.929** |
| | (2249.189) | (120.718) |
| (Expenses) Soap/Detergent/Cosmetics | 453.215 | 24.470 |
| | (840.648) | (49.555) |
| (Expenses) Leisure | 159.483 | 139.482*** |
| | (620.955) | (42.846) |
| (Expenses) Other | 155.419 | -1.246 |
| | (1249.261) | (59.513) |
| Number of youth | | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). We control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. Standard deviations of the control group are in parentheses, below the control means. Robust standard errors are in parentheses, below the ITT estimates. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A13: Impact on well-being and self-esteem

| | Control Mean | Mafita |
|---|---|---|
| Well-being | - | 0.114** |
| | - | (0.049) |
| *Components of index* | | |
| Cantril's Ladder today | 2.261 | 0.336*** |
| | (1.426) | (0.069) |
| Felt happy in past month | 3.170 | 0.042 |
| | (0.607) | (0.030) |
| Felt calm in past month | 3.168 | 0.054* |
| | (0.632) | (0.030) |
| Not nervous in past month | 2.959 | -0.019 |
| | (0.656) | (0.032) |
| Not blue in past month | 3.045 | 0.039 |
| | (0.698) | (0.034) |
| Not inconsolable in past month | 3.237 | -0.017 |
| | (0.690) | (0.033) |
| Self-esteem | - | -0.032 |
| | - | (0.048) |
| *Components of index* | | |
| Can prevent bad things from happening to self | 2.157 | 0.032 |
| | (0.741) | (0.036) |
| Can solve problems by yourself | 2.411 | -0.059 |
| | (0.817) | (0.040) |
| Future depends mainly on self | 2.267 | 0.018 |
| | (0.848) | (0.041) |
| Not helpless to deal with problems | 2.758 | -0.041 |
| | (0.756) | (0.037) |
| Rarely felt exploited or cheated | 2.334 | -0.084* |
| | (0.979) | (0.046) |
| Shared problems with friends and family | 3.285 | 0.039 |
| | (0.672) | (0.031) |
| Can achieve anything if dedicated | 3.601 | -0.002 |
| | (0.538) | (0.027) |
| Number of youth | | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). Each row presents the results of a separate regression The variables "Well-being" and "Self-esteem" are mean effect indices, standardized using the control group mean and standard deviation. The other variables are the individual outcomes that constitute these indices. We control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. Standard deviations of the control group are in parentheses, below the control means. Robust standard errors are in parentheses, below the ITT estimates. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table A14: Impact on expanded gender norms index (female youth)

| | Mafita | N |
|---|---|---|
| Gender norms index (youth) | 0.154** | 843 |
| | (0.061) | |
| *Components of index* | | |
| Marriage and fertility | 0.047 | 843 |
| | (0.060) | |
| Household decision-making | 0.175*** | 843 |
| | (0.062) | |
| Role in society | 0.131** | 843 |
| | (0.064) | |
| Domestic violence | -0.037 | 843 |
| | (0.068) | |
| Raven's matrix | 0.064 | 843 |
| | (0.070) | |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). The variable in the top panel is a mean effect index comprised of all gender norms outcomes. The bottom panel presents component mean effect indices that aggregate gender norms outcomes related to marriage and fertility, household decision-making, women's role in society, and domestic violence separately. In addition, the final row of the bottom panel presents the outcome from a Raven's matrix cognitive test. We control for strata fixed effects and demographics at randomization. Demographics include age, ethnicity, religion, and type of vulnerable group. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A15: Impact on caregiver gender norms index by gender of wards

| | Mafita |
|---|---|
| *Male wards* | |
| Gender norms index (caregivers) | 0.056 |
| | (0.128) |
| Female work decisions | 0.019 |
| | (0.121) |
| Marriage and fertility | -0.038 |
| | (0.114) |
| Role in society | 0.054 |
| | (0.135) |
| Household decision-making | 0.036 |
| | (0.135) |
| Number of caregivers | 268 |
| *Female wards* | |
| Gender norms index (caregivers) | 0.211** |
| | (0.091) |
| Female work decisions | 0.160* |
| | (0.093) |
| Marriage and fertility | 0.031 |
| | (0.078) |
| Role in society | 0.168* |
| | (0.089) |
| Household decision-making | 0.038 |
| | (0.089) |
| Number of caregivers | 518 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). The first variable in each panel is a mean effect index comprised of all caregiver gender norm outcomes. The other variables are component mean effect indices that aggregate gender norms outcomes related to female work decisions, marriage and fertility, women's role in society, and household decision-making separately. We control for strata fixed effects. Robust standard errors are in parentheses. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table A16: Impact on caregiver gender norms index, including caregivers with multiple wards

| | Mafita |
|---|---|
| Gender norms index (caregivers) | 0.123* |
| | (0.068) |
| *Components of index* | |
| Female work decisions | 0.062 |
| | (0.067) |
| Marriage and fertility | -0.007 |
| | (0.061) |
| Role in society | 0.146** |
| | (0.065) |
| Household decision-making | 0.027 |
| | (0.066) |
| Number of caregivers | 903 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). The variable in the top panel is a mean effect index comprised of all caregiver gender norm outcomes. The bottom panel presents component mean effect indices that aggregate gender norms outcomes related to female work decisions, marriage and fertility, women's role in society, and household decision-making separately. We control for strata fixed effects. Robust standard errors are in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A17: Impact on existing social ties

| | Control Mean | Mafita |
|---|---|---|
| Live with Mallam | 0.169 | -0.004 |
| | (0.375) | (0.013) |
| Currently studying with Mallam | 0.634 | -0.006 |
| | (0.482) | (0.019) |
| Confidence in religious leaders of own religion | 3.857 | 0.037** |
| | (0.389) | (0.017) |
| Trust in your family | 3.954 | 0.003 |
| | (0.231) | (0.012) |
| Trust in people in neighborhood | 3.355 | -0.011 |
| | (0.684) | (0.032) |
| Trust in people you know personally | 3.464 | -0.020 |
| | (0.698) | (0.032) |
| Trust in people you meet for the first time | 2.409 | 0.064 |
| | (0.851) | (0.040) |
| Trust in people from other ethnic group | 2.449 | 0.077** |
| | (0.810) | (0.037) |
| Trust in people from other nationality | 2.305 | 0.111*** |
| | (0.827) | (0.039) |
| Trust in people from other state | 2.534 | 0.131*** |
| | (0.754) | (0.036) |
| Time spent with family (hours) | 10.991 | -0.048 |
| | (11.469) | (0.528) |
| Time spent socializing with friends (hours) | 5.950 | 0.817** |
| | (6.967) | (0.344) |
| Number of friends in religious school | 3.421 | 0.129** |
| | (1.843) | (0.058) |
| Involved in church/mosque | 0.300 | 0.031 |
| | (0.457) | (0.020) |
| Involved in youth group | 0.194 | 0.030* |
| | (0.393) | (0.018) |
| Involved in sports club | 0.067 | 0.025** |
| | (0.249) | (0.012) |
| Borrowed money from family/friends | 0.255 | -0.010 |
| | (0.436) | (0.021) |
| Number of youth | | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). We control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. Standard deviations of the control group are in parentheses, below the control means. Robust standard errors are in parentheses, below the ITT estimates. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

Table A18: Impact on religiosity, trust in other religions, and donations

| | Control Mean | Mafita |
|---|---|---|
| Religiosity | - | -0.085* |
| | - | (0.046) |
| *Components of index* | | |
| Time spent on religious activities (hours) | 20.656 | -0.606 |
| | (13.320) | (0.613) |
| Important : Religion (disagree) | 1.000 | -0.001 |
| | (0.000) | (0.001) |
| Other things more important than religion (disagree) | 0.733 | -0.041* |
| | (0.442) | (0.022) |
| Strong sense of God presence | 0.978 | -0.003 |
| | (0.148) | (0.007) |
| Trust in other religions | - | 0.146*** |
| | - | (0.046) |
| *Components of index* | | |
| Number of friends from different religion | 0.028 | 0.023 |
| | (0.230) | (0.014) |
| Trust in people from other religion | 2.362 | 0.099** |
| | (0.887) | (0.040) |
| Confidence in religious leaders of other religions | 2.486 | 0.105** |
| | (0.971) | (0.046) |
| Donation experiment | - | 0.017 |
| | - | (0.084) |
| | | [0.568] |
| Number of youth | | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). The table presents mean effect indices comprised of all outcomes related to religiosity and trust in other religions, respectively, as well as the results of a donation experiment intended to measure generosity towards other religions. The variables in the "components of index" panels are the component variables that make up the index in the panel above them. We control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. Standard deviations of the control group are in parentheses, below the control means. Robust standard errors are in parentheses, below the ITT estimates. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A19: Impact on religious enforcement and anti-social behavior

| | Control Mean | Mafita |
|---|---|---|
| Religious enforcement | - | 0.033 |
| | - | (0.044) |
| | | [0.502] |
| *Components of index* | | |
| Do more to ensure people follow religious laws | 0.399 | 0.003 |
| | (0.490) | (0.023) |
| Would punish: Alcohol | 0.075 | 0.020 |
| | (0.263) | (0.013) |
| Would punish: Mini skirt | 0.085 | -0.010 |
| | (0.279) | (0.013) |
| Would punish: Blasphemy | 0.231 | 0.015 |
| | (0.421) | (0.020) |
| Anti-social behaviour | - | 0.064 |
| | - | (0.049) |
| | | [0.236] |
| *Components of index* | | |
| Taken Tramadol or Beneline | 0.018 | 0.011 |
| | (0.132) | (0.008) |
| Taken kayan maye | 0.048 | 0.007 |
| | (0.212) | (0.011) |
| Stolen when victim was not present | 0.008 | 0.010* |
| | (0.091) | (0.006) |
| Stolen when victim was present | 0.017 | 0.003 |
| | (0.128) | (0.007) |
| Done some work for a criminal group | 0.008 | 0.007 |
| | (0.091) | (0.005) |
| Gotten into a physical fight | 0.036 | -0.013 |
| | (0.185) | (0.008) |
| Number of youth | | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). The table presents mean effect indices comprised of all outcomes related to religious enforcement and anti-social behavior, respectively. The variables in the "components of index" panels are the component variables that make up the index in the panel above them. We control for strata fixed effects. Robust standard errors are in parentheses. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A20: Impact on political and religious violence

| | Control Mean | Mafita |
|---|---|---|
| Political and religious violence | - | 0.029 |
| | - | (0.048) |
| | | [0.502] |
| *Components of index* | | |
| Used force or violence for political cause | 0.025 | -0.003 |
| | (0.156) | (0.007) |
| Took part in group admin of justice | 0.060 | 0.005 |
| | (0.238) | (0.012) |
| Used force or violence for religious reason | 0.151 | 0.033* |
| | (0.357) | (0.018) |
| Participated in riot | 0.017 | -0.003 |
| | (0.128) | (0.006) |
| Attitudes toward political violence | - | -0.004 |
| | - | (0.049) |
| *Components of index* | | |
| Use of violence is justified in protesting an injustice | 0.164 | -0.023 |
| | (0.371) | (0.018) |
| It is good to use violence to resolve problems | 0.078 | 0.004 |
| | (0.268) | (0.013) |
| It is ok to use force or violence for political cause | 0.077 | 0.009 |
| | (0.266) | (0.013) |
| Number of youth | | 1706 |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). The outcome variable in the top panel is a mean effect index, standardized using the control group mean and standard deviation. We control for strata fixed effects. Robust standard errors are in parentheses. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table A21: Heterogeneity by classroom gender composition

| | Mafita | Mafita × Gender mix | |
|---|---|---|---|
| *Primary* | | | |
| Income and employment | 0.239*** | 0.001 | 1706 |
| | (0.047) | (0.002) | |
| | [0.001] | | |
| Gender norms index (youth) | 0.118*** | 0.002 | 1706 |
| | (0.044) | (0.002) | |
| | [0.022] | | |
| Gender norms index (caregivers) | 0.160** | 0.002 | 792 |
| | (0.071) | (0.003) | |
| | [0.045] | | |
| *Secondary* | | | |
| Assets and expenditures | 0.166*** | 0.002 | 1706 |
| | (0.048) | (0.002) | |
| Job search behaviour | 0.086* | 0.001 | 1706 |
| | (0.048) | (0.002) | |

*Notes:* This table reports estimated effects of assignment to treatment, interacted with a measure of the gender composition of the classroom. This gender mix variable is defined as the absolute value of the percentage of girls minus the percentage of boys in the trade and center the participant applied to. We de-mean this variable by its sample mean, so the uninteracted treatment coefficient can be interpreted as the effect of treatment at the average gender mix in the program. The mean and standard deviation of the gender mix variable are 0.33 and 0.25, respectively. All outcome variables are mean effect indices, standardized using the control group mean and standard deviation. Each row presents results from a separate regression. All regressions, except for the caregiver gender norms index, control for individual characteristics at the time of randomization (age, gender, ethnicity, religion, and type of vulnerable group) and strata fixed effects. Robust standard errors are in parentheses. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A22: Impact on economic outcomes and gender norms, excluding youth selected for additional entrepreneurship training

| | Mafita | N |
|---|---|---|
| *Primary* | | |
| Income and employment | 0.218*** | 1335 |
| | (0.057) | |
| | [0.002] | |
| Gender norms index (youth) | 0.129** | 1335 |
| | (0.052) | |
| | [0.040] | |
| Gender norms index (caregivers) | 0.163** | 792 |
| | (0.073) | |
| | [0.048] | |
| *Secondary* | | |
| Assets and expenditures | 0.171*** | 1335 |
| | (0.055) | |
| Job search behaviour | 0.075 | 1335 |
| | (0.058) | |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). All outcome variables are mean effect indices, standardized using the control group mean and standard deviation. For the youth sample, we control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. For the caregiver sample, we control for strata fixed effects. Robust standard errors are in parentheses. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. The sample is restricted to youth who were not selected for additional entrepreneurship training. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A23: Impact on economic outcomes and gender norms, using ICW indices

| | Mafita | N |
|---|---|---|
| *Primary* | | |
| Income and employment | 0.259*** | 1706 |
| | (0.049) | |
| | [0.001] | |
| Gender norms index (youth) | 0.116*** | 1706 |
| | (0.044) | |
| | [0.026] | |
| Gender norms index (caregivers) | 0.167** | 792 |
| | (0.071) | |
| | [0.032] | |
| *Secondary* | | |
| Assets and expenditures | 0.188*** | 1706 |
| | (0.046) | |
| Job search behaviour | 0.134*** | 1706 |
| | (0.048) | |

*Notes:* This table reports estimated effects of assignment to treatment (ITT). All outcome variables are inverse-covariance weighted indices, standardized using the control group variance. For the youth sample, we control for strata fixed effects and demographics at randomization. Demographics include age, gender, ethnicity, religion, and type of vulnerable group. For the caregiver sample, we control for strata fixed effects. Robust standard errors are in parentheses. Multiple inference-adjusted p-values, adjusted across all primary outcomes, are in brackets. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.