Gender Norms and the Division of Labor within the Household *

Giang Thai †

December 16, 2025

Abstract

Can gender training interventions shift household labor allocation and women's decision-making power? I address this question using a randomized control trial of a multi-year gender training program in Bangladesh that targeted women and encouraged men's involvement. On average, I find little to no significant long-term effects on household division of labor or women's decision-making power. However, combining rich women's empowerment data and machine learning techniques, I find substantial heterogeneity in treatment effects by community gender norms. For the most affected households associated with communities having relatively less restrictive norms, women enhance their participation in income-generating and consumption decisions while men engage more in household chores. In contrast, among the least affected households associated with relatively more conservative areas, women experience reduced time in income-generating work and decreased decision-making power. These findings suggest that gender training interventions may backfire in highly conservative settings, highlighting the importance of accounting for community norms when designing and targeting gender interventions.

Keywords: Gender, Intra-household Allocation, Labor Supply, Field Experiments
JEL Codes: J16, D13, J22, C93

*I am grateful for the IFPRI team that collected the data for the TRAIN program, led by Marie Ruel, Neha Kumar, Agnes Quisumbing, and John Hoddinott. I thank Neha Kumar and Agnes Quisumbing for generously sharing the data with me for this paper and for advice through many helpful conversations about the analysis. I benefited greatly from feedback provided by Marc Bellemare, Jason Kerwin, Paul Glewwe, Elizabeth Davis, Steve Polasky, Caroline Krafft, and my colleagues in APEC 8903-8904. The opinions expressed here belong to the author and do not necessarily reflect those of IFPRI or UMN.

†University of Minnesota. Email: thaix034@umn.edu.

1 Introduction

Across countries, the burden of housekeeping and childcare falls disproportionately on women. The implication of this is that women have fewer hours and less flexibility after fulfilling these responsibilities, limiting their ability to work outside the household. The resulting constraints on women's labor force participation lead to lower lifetime earnings, reduced control over household resources, and diminished bargaining power in household decisions for women. Globally, women's labor force participation lags behind men's, with particularly striking gaps of approximately 55 and 50 percentage points in the Middle East and North Africa and South Asia respectively (Heath et al., 2025). This misallocation of labor, often reinforced by entrenched gender norms, not only perpetuates inequality but also represents a significant economic inefficiency, limiting both household welfare and aggregate productivity.

Common approaches to increasing female employment rates include interventions aimed at freeing women's time from household responsibilities, such as introducing labor-saving technologies (Bharati, Qian and Yun, 2021) and subsidizing childcare (de Barros et al., 2011). An alternative approach with more limited evidence is to directly shift gender attitudes and norms regarding gender roles in household responsibilities. While several programs incorporate gender training as an add-on component, rigorous evidence on whether such trainings actually translate into behavioral changes in household labor allocation, rather than merely shifting attitudes, remains scarce.

I study whether participating in a gender training intervention leads to changes in the household division of labor, particularly shifting household responsibilities between women and men. Specifically, I examine whether gender training that engages both partners leads men to increase their participation in domestic activities while enabling women to increase their participation in income-generating activities (IGAs). The hypothesized mechanism is that training can shift internalized beliefs about appropriate gender roles, which in turn affects household bargaining and the willingness of both partners to renegotiate the division of labor. Understanding whether this mechanism operates in practice has important policy implications. Unlike labor-saving technologies or childcare subsidies that require costly infrastructure investment, gender trainings, if effective, could

offer a scalable and cost-effective intervention that generates persistent changes in household behavior, reaching women in settings where infrastructure investments are infeasible or prohibitively expensive.

I identify the causal impact of gender trainings using data from a randomized control trial (RCT) in Bangladesh implemented by BRAC, a large non-governmental organization. Bangladesh provides a particularly informative setting to test whether gender training can shift entrenched behaviors. Specifically, gender norms around household roles are especially rigid, and women's labor force participation remains stagnantly low despite the country's significant progress in women's education and microfinance access. The intervention, part of BRAC's Targeting and Re-aligning Agriculture to Improve Nutrition (TRAIN) program, combined household-level activities (monthly home visits) with community-level engagement (bimonthly gender forums with couples and community members). Importantly, while the program primarily targeted women, men's participation was strongly encouraged, allowing me to examine whether male engagement enhances the intervention's impact on household labor reallocation. I measure impacts using data collected one and a half years after the program ended, thus capturing long-run behavioral changes rather than temporary responses to the intervention.

I find that, on average, gender trainings have little to no impact on the long-run household division of labor. The finding is consistent for both time use and decision-making outcomes. Women who received gender trainings are no more likely to participate in paid work and income-generating decisions than their counterparts in the control group. The finding also applies to men's participation in domestic work. This lack of movement in the outcomes of interest indicates that the status quo of gender roles within the household remains unchanged. While it is possible that the program had short-run effects on household labor division, due to data limitations, I cannot test this hypothesis in this study. Results also show that both women and men's engagement in the gender training makes no difference in changing the household's norms around task division. Men's participation was relatively low at 36 percent compared to women's at 60 percent. This suggests that, in this context, men's engagement and exposure to gender-equitable messages are

insufficient to change the status quo in the long run.

I rule out the possibility that the main findings were driven by ineffective implementation of the gender training. Despite suggestive evidence that control group members received the training as well, when examining differential levels of exposure to gender messages by treatment status, I find that both treated men and women could recall a significantly higher number of gender topics discussed in the training and were more likely to recall discussing topics related to division of labor within the household, with the control group recall at relatively low levels. I estimate local average treatment effects to address the potentially endogenous participation in the gender training.

A possible reason for the null impact of the gender training is entrenched social norms regarding gender roles within households. One's behaviors depend not only on their own preferences but also on other people's behaviors and expectations. Studies on social norms show that it is difficult to change individual behavioral patterns when other people in the community hold on to entrenched behaviors; change can occur when there is a tipping point where a sufficient number of people are willing to change their behaviors (Nyborg et al., 2016). While I do not have direct measures of gender norms through first- and second-order beliefs, I take advantage of rich women's empowerment data at baseline. These data were designed to capture multi-dimensional aspects of women's empowerment, ranging from input into productive decisions to access to credit to asset ownership. I combine aggregated women's empowerment indicators at the community level, serving as proxies for gender norms, with machine learning techniques following the procedure proposed by Chernozhukov et al. (2025) to examine heterogeneous treatment effects across four quartiles, with the top quartile representing households most affected and the bottom quartile representing households least affected by gender training.

I find that positive impacts for most affected households but negative impacts for least affected households. Treated women in most affected households are significantly more likely to engage in paid work or IGA decisions, while treated men are significantly more likely to participate in domestic work. But the results reverse for least affected households. In particular, treated women are *less* likely to participate in IGAs and have *lower* decision-making power. The differences in

outcomes between the top and the bottom quartiles are statistically significant. When examining baseline values of women's empowerment indicators of the top and bottom quartiles of treatment effects, I find that for the outcomes on women's decision-making and men's time on chores, most affected households come from communities with higher female non-farm employment, more inputs in productive decisions, and higher access to credit. For the outcome on women's time on productive work, the most affected households come from communities in which women have *lower* asset ownership, less mobility, and are less likely to be empowered. These findings point toward a direction that the gender training intervention is more effective in shifting some margins of gender roles in communities that already have relatively less restrictive gender norms, nudging them toward more equitable gender roles. But they also emphasize a word of caution for implementing similar interventions in places with highly entrenched norms where backlash could lead to worsened outcomes for women.

This paper contributes to the growing research on information-based interventions intended to increase women's empowerment in developing countries. Various studies that have a gender training add-on through discussion groups have shown mixed impacts on gender attitudes and decision-making dynamics. The IMAGE program in South Africa finds that neither microfinance nor microfinance combined with training on gender and HIV results in impacts on women's decision-making autonomy (Kim et al., 2009). Another savings group program in Côte d'Ivoire with a gender dialogue intervention shows changes in gender attitudes only among participants with high attendance rates (Gupta et al., 2013). Meanwhile, many training programs target adolescents and young adults using schools or community engagement as the main platform of implementation (Buchmann et al., 2018; Bandiera et al., 2020; Dhar, Jain and Jayachandran, 2022; Ashraf et al., 2020). Buchmann et al. (2018) find that empowerment training and financial incentives to delay marriage in Bangladesh do not decrease child marriage, but they do increase the probability that eligible girls stay in school. Bandiera et al. (2020) show that adolescent girls in communities receiving female empowerment and livelihood training in Uganda are more likely to be self-employed. Dhar, Jain and Jayachandran (2022) find positive impacts of classroom discussions about gender equality,

shifting secondary school boys and girls toward more gender-equity attitudes in India.

Within this literature, this paper makes a further contribution by focusing on an information-based gender training intervention by itself that targets adult married women and encourages men's engagement. The null effects of the program, even for households in which both the woman and the man participated in the training, shed light on the challenges in changing the gendered division of labor when the norms are deeply entrenched in daily life activities and widespread in the community. It reveals that, in this context, while possibly important, men's engagement is insufficient to change household norms.

This paper also contributes to research on constraints to female labor force participation. Studies in developing countries have explored different ways to influence attitudes toward female employment and free up women's time. Dean and Jayachandran (2019) evaluate the impact of showing a "family-orientation" video that addressed common concerns about safety in women's work among family members in India and do not find a measurable impact. Talamas Marcos (2023) finds that the death of grandmothers in Mexico leads to a sharp fall in the mother's employment rate, highlighting the importance of the availability of childcare on female employment. McKelway (2025) for the context of India, demonstrates that showing family members a promotional video about job opportunities for women in carpet weaving leads to a large increase in short-run female employment but it comes at the expense of women's leisure time. My contribution is to provide further evidence on the effectiveness, or lack thereof, of a gender training intervention by itself on the long-run household division of labor. My findings are consistent with the null impacts of other interventions on long-run female employment, thus highlighting the challenge in shifting intra-household allocation of tasks and decision-making dynamics.

The paper follows the following road map: Section 2 describes the TRAIN program and its experimental design. Section 3 provides information about the data, outcomes, and power calculations. Section 4 presents balance tests and issues with selection. Section 5 and Section 6 detail the empirical framework and average effects respectively. Section 7 discusses heterogeneous effects and Section 8 concludes.

2 Intervention and Experimental Design

2.1 Gender Trainings: Context and Program Details

The paper evaluates the impact of the gender sensitization component as part of the TRAIN program built on BRAC's standard micro-credit program, or so-called DABI. Households in the TRAIN program were selected based on three criteria, including (i) at least one female household member must have taken a DABI loan in the last 12 months; (ii) household must have had at least three years' experience in farming (crops, vegetables, fisheries or livestock); and (iii) either the DABI client was married and in between 15 and 28 years old at baseline, or there was at least one married woman between the ages of 15 and 28 residing in the household. The selected woman between 15 and 28 years old at baseline in each household is the index woman, and her husband (if co-living) is the index husband. The TRAIN program consists of four components. The first component is BRAC's business-as-usual DABI activities, targeted to low-income women to enable them to buy productive assets, invest in business, or smooth consumption and manage shocks. The second component is a nutrition behavior change communication (BCC) intervention. It consisted of once-per-month household visits by female frontline workers to provide household members with better nutrition knowledge for women and children under two years old. The third component is the nutrition-sensitive agriculture extension services with three main activities, which included home visits by agricultural frontline workers to provide advice and practical support on agricultural production, development of homestead gardens to ensure adequate production of nutritious foods, and additional training sessions in BRAC subdistrict offices.

The fourth component is the gender sensitization component, also referred to as the gender training. The intervention aims to sensitize husbands and other household members and raise their awareness of various women and girls' issues and of more gender-equitable household responsibilities. This component has three main activities. The first is an additional monthly home visit. The second is gender forums aimed to raise awareness of gender issues and seek commitment to action by families and influential community members. The forums were organized once every two

months for four or five beneficiary households, including index couples and their family members. The last activity was additional gender trainings for index couples at the BRAC regional office. Below are the 12 topics covered in the training sessions at both the household and community levels.

1. Workload, rest and recreation
2. Women's participation in income generating activities
3. Access to and control over resources/assets
4. Sexual harassment and dowry
5. Child marriage and early pregnancy
6. Violence against women and children
7. Women's mobility and involvement in household decision-making
8. Governmental and non-governmental services
9. Environment and women
10. Health and nutrition
11. Education
12. Health service and treatment

The intervention started in March 2017 and was planned to continue until April 2020. However, in March 2020, at the onset of the COVID-19 pandemic, the implementation activities were suspended until June 2020. Then the program was resumed with pandemic protocols, and officially ended on October 31st, 2020. The intervention was implemented in 144 rural unions (the smallest administrative unit in Bangladesh) in ten districts across three divisions – Rangpur, Dhaka and Khulna – in Bangladesh.

2.2 Experimental Design

The study is a cluster randomized control trial (RCT) with one control group and three treatment arms in an additive design, in which unions are clusters and divisions are strata. The experiment design is illustrated in Figure I. Within each of the three divisions, 48 unions were randomly

selected to be in four groups described as follows:

- Treatment arm 1 or Control group or Base DABI: The group is the business-as-usual group, in which beneficiaries received BRAC's standard micro-credit program.
- Treatment arm 2 or Base DABI + Nutrition BCC: The group received BRAC's standard micro-credit program plus the nutrition BCC component.
- Treatment arm 3 or Base DABI + Nutrition BCC + Agricultural extension: The group received BRAC's standard micro-credit program plus the nutrition BCC component as well as the extension services.
- Treatment arm 4 or Base DABI + Nutrition BCC + Agricultural extension + Gender sensitization: The group received BRAC's standard micro-credit program, the nutrition BCC component, the extension services, plus the gender sensitization component.

The TRAIN program was pre-registered on the 3ie Registry for International Development Impact (Ruel and Kumar, 2023) prior to the data analysis phase. However, due to a different research question that is not part of the main impact evaluation, this paper lies outside of the scope of the pre-analysis plan, thus not following the pre-specified analysis closely.

For this paper, to isolate the impact of the gender sensitization component, I categorize households in treatment arm 4 as the treated sample and households in treatment arm 3 as the control sample.

3 Data and Power Calculations

3.1 Data

The data come from the baseline and endline survey data, collected separately for women and men. For women, the primary respondent in each household was the selected index woman. For men, the primary respondent was the husband of the index woman. If the index husband either had limited involvement in household agricultural activities or was absent, another adult male who was most knowledgeable about agricultural activities in the household was interviewed instead.

The baseline survey took place from November 2016 to January 2017 for 5,040 households. Due to the COVID-19 pandemic, the endline data collection was delayed and rescheduled to March 2022. Since the paper examines the gendered division of labor, the sample chosen for the analysis is restricted to dual-adult households found at endline. I discuss in detail the balance tests and selection into dual-adult households in Section 4.

To estimate changes in the division of labor within the household, I examine two main types of outcomes, time use and decision making. For the time use data, I am interested in the share of awake time that the woman and man spent on IGAs and domestic work. I list out the activities categorized under IGAs and domestic work in the Appendix Table 1. For the decision making data, I use the share of decisions related to income-generating activities (IGA) and household consumption.

At baseline, unsurprisingly, women tend to spend more time on the domestic work category, such as cooking, domestic work, and caring for children. With regards to the income-generating work category, women tend to do more poultry or small animal raising. On the other hand, men are likely to spend more time on the income-generating work and mobility categories. The former type includes employed work, working on a household business, and staple grain farming, while the latter entails commuting, shopping or getting services, and traveling (not for work reasons). Figure II provides examples of the task division across households in the study sample based on the time use data at baseline where the division of labor reflects clearly the traditional gender roles within a household. Similarly, the activities in which women dominated the decision-making process at baseline include poultry and other small animals raising, as well as routine household purchases. Meanwhile, men dominated decisions related to staple grain farming, non-farm economic activities, wage and salary employment, and large occasional household purchases.

3.2 Power Calculations

As mentioned above, Treatment arm 3 serves as the control group since it has all the same components as treatment arm 4, except for the gender training, thus making it the cleanest control group.

I conduct ex-post power calculations for all outcomes with a significance level of 0.05 to detect increases of 20, 30, and 40 percentage points in the outcomes, adapting the practice in (McKenzie and Woodruff, 2014). Table I shows the power needed for the comparison between treatment 3 and treatment 4. The power of 80 percent or higher is desirable in field experiments and medical trials (Schulz and Grimes, 2005; Duflo, Glennerster and Kremer, 2007).

While I could not find any studies in the literature that have the same outcomes used in this paper, I rely on studies that aimed at increasing female employment in developing countries through information-based interventions as the benchmark for how large the effect should be. McKelway (2025) reports a 32.8 percentage increase in off-farm female employment and a 62.8 percent increase in women's time on paid work by showing family members a promotional video of female employment. Jensen (2012) finds that providing women with information about jobs increases female paid employment outside the home by 11.4 percentage points among the treatment group compared to the control group. Table I shows that I have adequate power to detect a 30 percentage increase for two out of four main outcomes and a 40 percentage increase for three out of four main outcomes when the control group is treatment arm 3.

4 Balance and Selection

4.1 Balance

I present summary statistics of baseline characteristics of households and index couples and balance tests between the treatment and control groups in Tables!II and III for two samples, (i) the original study sample comprised of 2,520 households in Treatments 3 and 4, and (ii) the dual-adult household sample at endline with 2,150 households. Definitions of a few baseline variables that are not self-explanatory are shown in the Appendix Table 1.

Overall, index women were 23 years old on average at baseline, had approximately six years of education, were married at 16 years old, and had an average of 1.4 children. Around 86 percent of them co-resided with their husbands. Seventy percent of the women were DABI beneficiaries. For

those living with their husbands, their husbands were 30 years old, with five years of schooling, on average. Households in the study had close to five members and their consumption expenditures were 3,210 Bangladesh Taka (BDT) per month per capita on average.

In the original study sample, at the ten percent level, treatment arm 4 households are significantly different from treatment arm 3 for seven characteristics: female household head, ratio of male to female household members, dependency ratio, dwelling with electricity, index woman's age, her years of schooling, husband's co-residence. The omnibus F-tests of joint orthogonality with randomization inference, following Kerwin, Rostom and Sterck (2024), is insignificant (p = 0.41). For dual-adult households, the means in treatment arm 4 statistically differ from the ones in treatment arm 3 for six variables: female household head, ratio of male to female household members, dependency ratio, dwelling with electricity, index woman's age, and husband's co-residence. The omnibus F-test is insignificant as well (p = 0.79). Thus, I have a balanced sample, regardless of restricting to dual-adult households or not.

4.2 Selection into Non-Attrition and Dual-adult Households

Next, I examine the attrition rate and whether it differs across treatment arms in Tables IV and V. Considering the original study sample, the attrition is approximately five percent, as shown in Table IV. But when limiting to dual-adult household sample, the attrition rate is at ten percent.

Although the treatment assignment does not predict attrition, baseline characteristics of households and index women do predict the probability of having dual adults, shown in columns (1) and (2) in Table V. Non-attrited treatment households had a higher ratio of male to female members, the index woman had slightly more years of schooling, and with more children. The joint F-test using the randomization inference procedure is 0.004. Thus, I adjust for this potential selective attrition bias by creating inverse probability weights of being a non-attrited household and including them in the estimation model.

I also check whether there is selection into being dual-adult households at endline, shown in columns (3) and (4) of Table V. The selection issue is very similar to the selective attrition bias

mentioned above. Dual-adult households also had a higher ratio of male to female members, the index woman had more years of schooling, and with more children. The joint F-test using the randomization inference is 0.000. Therefore, I also create inverse probability weights of being a dual-adult household and multiply them with the attrition weights to account for these two selection issues.

4.3 Intervention Participation

The participation rates in gender training in the treatment group were relatively low. As Table VI indicates, 59.8 percent of women and 36.2 percent of men in treatment households reported participating in at least one gender training sessions between March 2019 and March 2020. It is not surprising to see men's lower engagement level compared to women's since the main program targets women while men's attendance is encouraged and not compulsory. While it is possible that gender training sessions took place when men were busy working outside the household, men's limited participation in the gender training sessions suggests their low interest in learning about gender equitable messages. Together, for 30 percent of households, both the woman and man attended the gender training during this time period.

The participation rates in gender training in the control group was not as low as one would expect if the program had been implemented according to the research design. In particular, in treatment arm 3, 34.8 percent of women and 16.2 percent of men reported their participation in gender training sessions. If the implementation had followed the research design, so that all households followed the random assignment, this rate should have been close to zero. It is possible that because treatment arm 3 households experienced various types of training (nutrition plus agricultural extension), they were confused when asked about gender training, thus misreporting their answers.

To provide further evidence on whether the high participation rate among treatment arms 1-3 households is due to respondents' confusion or due to the program implementation issue, I examine the impact of the gender training on exposure to gender equitable messages using the original

treatment assignment (T1-T4). If it were due to respondents' confusion during data collection, treatment arm 3 respondents would have had no differential exposure to the gender training content compared to their treatment arm 1 or 2 counterparts. However, as Table VII indicates, I find that treatment arm 3 households were significantly more likely to recall topics related to the division of labor in gender training sessions than treatment arms 1 and 2 households. They also recognized and discussed significantly more gender-related topics with others in the village. As a result, I cannot rule out a possible contamination bias in the program implementation. The presence of the contamination and attrition results in the issue of two-way non-compliance, in which there are people in the treatment group who did not take up the program and there are people in the control group who took up the program when they were not assigned to treatment. The next section on the empirical framework proposes an estimation strategy to maintain the internal validity of the impact results in the presence of the two-way non-compliance issue.

5 Empirical Framework

5.1 Estimation Strategy

I estimate the Local Average Treatment Effect (LATE) of the gender training intervention on the gendered division of labor. I use the instrument variable (IV) method to evaluate the impact of gender training on a specific group of participants. Assuming that those in the control group who took up gender trainings were always-takers, meaning that they would always take up the intervention regardless of their treatment status, the chosen method takes advantage of the initial random assignment as the instrument variable. The LATE estimate measures the impact among the compliers, referring to those who were induced to take up the intervention due to their treatment assignment. The estimation model runs in two stages as follows:

$$Y_{ij}^{1} = \beta_0 + \beta_1 \hat{P}_{ij}^{1} + \delta \boldsymbol{X_{ij}} + \varepsilon_{ij} \tag{1}$$

$$P_{ij}^{1} = \gamma_0 + \gamma_1 T4_j + \sigma \boldsymbol{X_{ij}} + u_{ij} \tag{2}$$

where Y_{ij}^1 denotes the outcome for household i in union j at endline. $T4_j$ equals one if a union j was assigned to the gender sensitization component and zero otherwise, and $\boldsymbol{X_{ij}}$ are baseline control and other control variables, including baseline characteristics and division (strata) fixed effects. In addition, P_{ij}^1 is the endogenous participation variable, in which P_{ij}^1 equals one if both the index woman and the man reported their participation, and zero otherwise. The term β_1 in Equation (2) is the coefficient of interest that measures the impact of the gender training among the compliers.

Adding division fixed effects as strata dummy variables can help increase the precision of the estimates Bruhn and McKenzie (2009). For any observations with missing baseline controls, I impute missing values using the sample mean and add binary indicators for missing observations to the estimation (Zhao and Ding, 2022). Due to the treatment assignment at the union level, in which randomly selected households within the assigned union were offered the gender training, clustering standard errors at the level of the random assignment is relevant and recommended (Abadie et al., 2022). Thus, the standard errors are clustered at the union level.

5.2 Identification Strategy

The LATE estimate adjusts for the two-sided non-compliance issue by dividing the intent-to-treat (ITT) by the difference between the take-up of the treatment group and the one of the control group (Gerber and Green, 2012; Glennerster and Takavarasha, 2013). The IV model measures LATE consistently under four assumptions. The first is the exclusion restriction, in which conditional on control variables, the instrument variable cannot affect the outcome in any way other than through the endogenous variable. The second is the relevance condition, in which conditional on control variables, the instrumental variable must be correlated with the endogenous variable. The third is the Stable Unit Treatment Value Assumption (SUTVA), in which one's outcome value does not depend on others' treatment status. Last, but not least, is the monotonicity assumption, which is the assumption of no defiers; this means that if one's treatment status is affected by the instrument variable, the endogenous variable would go in the same direction (Imbens and Angrist, 1994).

I examine whether these four assumptions hold in this study. The instrumental variable is the random assignment to treatment arm 4 while the endogenous variable is the self-reported participation in the gender training. The exclusion restriction assumption means that the only way that the assignment to treatment arm 4 can affect the gendered division of labor is through the participation in gender training. Given that this is based on an RCT, this is a reasonable assumption to make. The relevance condition assumption indicates that the assignment to treatment arm 4 must be correlated with participation in gender training. This assumption is satisfied in this case as the effective F-statistics in the first-stage results, shown in Appendix Table 3, reject the null of weak instruments (Olea and Pflueger, 2013). For the SUTVA assumption, given random assignment is at the union level, it is unlikely that the impact on the outcome of the treatment group spills over into the control group. The monotonicity assumption cannot be directly tested. Thus, the IV estimate offers LATE assuming that there are no defiers in the sample.

6 Average Impacts: Local Average Treatment Estimates

First, I look at the impact of gender trainings on the household division of labor using time use data, shown in columns 1 and 2 of Table VIII. For households in which both the woman or the man attended gender training, women are 0.7 percentage points less likely to participate in productive work, which is a 5.6 percent decrease from the control group mean. Men are 3.1 percentage point more likely to spend time on domestic activities at endline compared to their counterparts in other arms, which is a 119 percent increase from the control group mean. Although the coefficient is statistically insignificant, the positive and large magnitude suggests a potential shift in men's household responsibilities.

The results using decision-making outcomes are consistent with the time-use patterns observed above, as shown in columns 3 and 4 of Table VIII. Treated women involve in less share of decisions related to income-generating activities as well consumption decisions. The magnitudes of the coefficients are small compared to the control group mean. This means that in households where

both the woman and the man reported their participation, women who received gender trainings are no more likely to gain more power in decision making than those without gender trainings, if not a bit less. Treated women participated in 7.9 percentage points less in productive decisions and 2.5 percentage points less in consumption decisions, equivalent to 18.9 percent and 9.7 decreases respectively from untreated women. All these coefficients are statistically insignificant as well, implying the null impact of gender trainings. Overall, given the small sizes and the insignificance of the coefficients, it is clear that the intervention has little to no impact on the decision-making process in household activities.

7 Heterogeneous Treatment Effects

While the literature in gender training suggests that men's engagement in the training sessions could be beneficial for women (Chang et al. 2020), the findings from this study imply that men's participation with women in the training does not make a significant difference in changing the long-run gendered division of labor. Despite of the lack of long-run effects on the household division of labor, it is possible that the program could have resulted in some short-run effects if the endline data were collection sooner. Due to the COVID-19 pandemic, the data collection was delayed significantly, leading to the lack of short-term data needed to examine this hypothesis.

The lack of effect of gender training on the long-run household norm of task sharing, even with male engagement, is consistent with a few studies in the literature. It could stem from three reasons. First, even when the man participated in gender training, he did not find these gender equitable messages convincing and was not interested in changing intra-household arrangements with his spouse. In this case, the man is the gatekeeper of change and there is only so much a woman can change when her husband strongly agrees with the traditional gender roles. Second, household norms of task division could stem from internal labor allocation based on the couple's comparative advantage. It could mean that women have comparative advantage in making decisions regarding household chores while men have comparative advantage in income-generating

activities. However, due to data limitations, testing these two reasons is beyond the scope of this paper.

Third, it is possible that the couple were open to change their attitudes toward traditional gender roles, but that they lived in a community with strong norms of gendered labor division within the household. When norms are part of social coordination within a group, it is difficult to change individual behavioral patterns when other people in the community hold on to entrenched behaviors. Norms can change when there is a tipping point in the network or community, the point in which a sufficient number of people adopt new behaviors (Nyborg et al., 2016). To provide suggestive evidence for this reason, I find a proxy for the stickiness of gender norms to examine the heterogeneous impact of gender trainings in the next section.

The entrenchment of gender norms could impede the uptake and effectiveness of gender trainings in the first place (Chang et al. 2020). While direct measures of gender norms are not collected, I take advantage of the rich Women's Empowerment in Agriculture Index (WEAI) data at baseline to examine the heterogeneous treatment effects by community gender norms, covering 12 empowerment aspects, including production decisions, asset ownerships, access to credit, income decisions, autonomy in income, group membership, influential group membership, workload, mobility, respect, violence, and self-efficacy. To do so, I aggregate 88 pro-WEAI indicator variables, plus female non-farm employment rate at the union level. I apply the procedure proposed by Chernozhukov et al. (2025), which allows for the detection of high-dimensional treatment effect heterogeneity in field experiments, for all four main outcomes of interest.

Specifically, I split the sample into half, one for a training dataset and another for a testing dataset. With the training set, I generate the expected outcome for households with a full set of community-level covariates Z if they were assigned to control group and predicted treated effect using machine learning (ML) methods. I apply the four ML options, including random forest, support vector machine, neural net, and gradient boosting, and choose the method with the highest prediction scores λ and $\bar{\lambda}$, which are goodness of fit measures. Table IX shows these metrics for each ML method for each outcome. For the share of women's time on productive work, share of

productive decisions women involved, and share of consumption decisions women made, support vector machine yields the highest scores, thus selected to be the main ML method to use onwards. For the share of men's time on domestic work, the selected method is gradient boosting.

The procedure estimates produce two key parameters of heterogeneity, one is β_2 to test the null hypothesis of no heterogeneity, and another is the sorted group average treatment effect (GATES) to get at treatment effects by quartiles. For the first parameter, Table X shows the estimates of β_2 to be significantly different from 0, with the p-value < 0.01 for the share of women's time on productive work, share of productive decisions women involved, and share of consumption decisions women made, and p-value < 0.05 for the share of men's time on domestic work.

When breaking down by quartile, presented in Table XI and Figure III, the treatment effects of the bottom quartile, referring to households least affected by gender training, are actually *negative*. Specifically, women in these households spend less time on productive work and involve in less productive and consumption decisions, and the coefficients are statistically significant. Men also spend less time on household chores although it is not significant. Meanwhile, the treatment effects of the top quartile, referring to households most affected by gender training, are all positive and significant. Women in these households experience more time spent on productive work and more decision-making power while men spend more time on domestic work. Notably, the differences in treatment effects between the top and the bottom quartiles are all statistically significant.

It is important to understand which community characteristics of norms determine the heterogeneity in treatment effects. Table XII presents the full list of 88 women's empowerment variables and their baseline values of the bottom and top quartiles. Overall, the share of men's time on domestic work, share of productive and consumption decisions women involved share similar baseline community patterns. The most affected households reside in areas with a significantly higher female employment rate, in which women involve in more productive decisions, better access to and use of credit, experience more respect from household members, and are more self-efficient, but have lower inputs into income decisions and autonomy in income. Meanwhile, the share of women's time in productive work are driven by different community norm patters. The most af-

fected households reside in areas with lower asset ownership, in which women have some inputs in income decisions and are members of fewer groups but experience less mobility, justify fewer domestic violence scenarios, and less likely to be empowered in general. Taken all this together, households in communities with relatively less restrictive gender norms at baseline tend to be most affected by gender training and therefore experience positive shifts toward more equitable gender roles, particularly in terms of women's decision-making power and men's engagement in household chores. While the heterogeneity associated with the baseline women's empowerment indicators cannot be interpreted as causal effects, it emphasizes the large variation in treatment effects by community norms surrounding treated households.

8 Conclusion and Discussion

In this paper, I examine the impact of gender training on the gendered division of labor within the household using data from a field experiment in Bangladesh. I provide evidence on the effectiveness, or lack thereof, of gender training on changing gender roles, especially with men's engagement in training sessions and offer suggestive findings that sticky gender norms could play a role in inhibiting the effect of gender training to begin with.

I find that, on average, the gender training intervention does not lead to change in household's status quo of task division in both time use and decision-making dynamics. However, in areas with relatively less restrictive gender norms, measured using the rich women's empowerment data at baseline and machine learning algorithms, women are more likely to have more decision-making power and men are more likely to do domestic work. However, the result is reverse for women's time on productive work, in which women in areas with relatively *more* restrictive gender norms are likely to engage in productive work. This suggests that gender equitable messages resonate more with participants when they are already not so entrenched in the traditional gender roles. The adverse impact for the least affected households implies that sticky community gender norms may be at play and hinder change to take place.

It is beyond the scope of this paper to identify why gendered division of labor is so entrenched and hard to change. It would be insightful for future research to disentangle the extent to which the gendered division of labor is conditional on social expectations and identify other alternative interventions in which male engagement in changing household norms is more effective.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2022. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics*, 138(1): 1–35.

Ashraf, Nava, Natalie Bau, Corinne Low, and Kathleen McGinn. 2020. "Negotiating a Better Future: How Interpersonal Skills Facilitate Intergenerational Investment." *The Quarterly Journal of Economics*, 135(2): 1095–1151.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. 2020. "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa." *American Economic Journal: Applied Economics*, 12(1): 210–259.

Bharati, Tushar, Yiwei Qian, and Jeonghwan Yun. 2021. "Fuelling the Engines of Liberation with Cleaner Cooking Fuel." 59.

Bruhn, Miriam, and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1(4): 200–232.

Buchmann, Nina, Erica Field, Rachel Glennerster, Shahana Nazneen, Svetlana Pimkina, and Iman Sen. 2018. "Power vs Money: Alternative Approaches to Reducing Child Marriage in Bangladesh, a Randomized Control Trial." J-PAL.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Iván Fernández-Val. 2025. "Fisher–Schultz Lecture: Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, With an Application to Immunization in India." *Econometrica*, 93(4): 1121–1164. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA19303.

Dean, Joshua T., and Seema Jayachandran. 2019. "Changing Family Attitudes to Promote Female Employment." *AEA Papers and Proceedings*, 109: 138–142. Publisher: American Economic Association.

de Barros, Ricardo Paes, Pedro Olinto, Trine Lunde, and Mirela Carvalho. 2011. "The Impact of Access to Free Childcare on Women's Labor Market Outcomes: Evidence from a Randomized Trial in Low-income Neighborhoods of Rio de Janeiro." World Bank Economists' Forum.

Dhar, Diva, Tarun Jain, and Seema Jayachandran. 2022. "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India." *The American Economic Review*, 112(3): 29.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Chapter 61 Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*. Vol. 4, , ed. T. Paul Schultz and John A. Strauss, 3895–3962. Elsevier.

Gerber, Alan S., and Donald P. Green. 2012. *Field experiments : design, analysis, and interpretation.* W.W. Norton.

Glennerster, Rachel, and Kudzai Takavarasha. 2013. *Running Randomized Evaluations: A Practical Guide.* Princeton University Press.

Gupta, Jhumka, Kathryn L. Falb, Heidi Lehmann, Denise Kpebo, Ziming Xuan, Mazeda Hossain, Cathy Zimmerman, Charlotte Watts, and Jeannie Annan. 2013. "Gender norms and economic empowerment intervention to reduce intimate partner violence against women in rural Côte d'Ivoire: a randomized controlled pilot study." *BMC International Health and Human Rights*, 13(1): 46.

Heath, Rachel, Arielle Bernhardt, Girija Borker, Anne Fitzpatrick, Madeline McKelway, Andreas Menzel, Teresa Molina, and Garima Sharma. 2025. "Female Labor Force Participation." VoxDev.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467–475. Publisher: [Wiley, Econometric Society].

Jensen, Robert. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *The Quarterly Journal of Economics*, 127(2): 753–792.

Kerwin, Jason, Nada Rostom, and Olivier Sterck. 2024. "Striking the Right Balance: Why Standard Balance Tests Over-Reject the Null, and How to Fix it." *SSRN Electronic Journal*.

Kim, Julia, Giulia Ferrari, Tanya Abramsky, Charlotte Watts, James Hargreaves, Linda Morison, Godfrey Phetla, John Porter, and Paul Pronyk. 2009. "Assessing the incremental effects of combining economic and health interventions: the IMAGE study in South Africa." *Bulletin of the World Health Organization*, 87(11): 824–832. Publisher: World Health Organization.

McKelway, Madeline. 2025. "How Does Women's Employment Affect Their Time Use? Evidence from a Randomized Encouragement Design in India."

McKenzie, D., and C. Woodruff. 2014. "What Are We Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *The World Bank Research Observer*, 29(1): 48–82.

Nyborg, Karine, John M. Anderies, Astrid Dannenberg, Therese Lindahl, Caroline Schill, Maja Schlüter, W. Neil Adger, Kenneth J. Arrow, Scott Barrett, Stephen Carpenter, F. Stuart Chapin, Anne-Sophie Crépin, Gretchen Daily, Paul Ehrlich, Carl Folke, Wander Jager, Nils Kautsky, Simon A. Levin, Ole Jacob Madsen, Stephen Polasky, Marten Scheffer, Brian Walker, Elke U. Weber, James Wilen, Anastasios Xepapadeas, and Aart De Zeeuw. 2016. "Social norms as solutions." *Science*, 354(6308): 42–43.

Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics*, 31(3): 358–369.

Ruel, Marie, and Neha Kumar. 2023. "RIDIE."

Schulz, Kenneth F, and David A Grimes. 2005. "Sample size calculations in randomised trials: mandatory and mystical." *The Lancet*, 365(9467): 1348–1353.

Talamas Marcos, Miguel Ángel. 2023. "Grandmothers and the gender gap in the Mexican labor market." *Journal of Development Economics*, 162: 103013.

Zhao, Anqi, and Peng Ding. 2022. "To Adjust or not to Adjust? Estimating the Average Treatment Effect in Randomized Experiments with Missing Covariates." *Journal of the American Statistical Association*, 1–11.

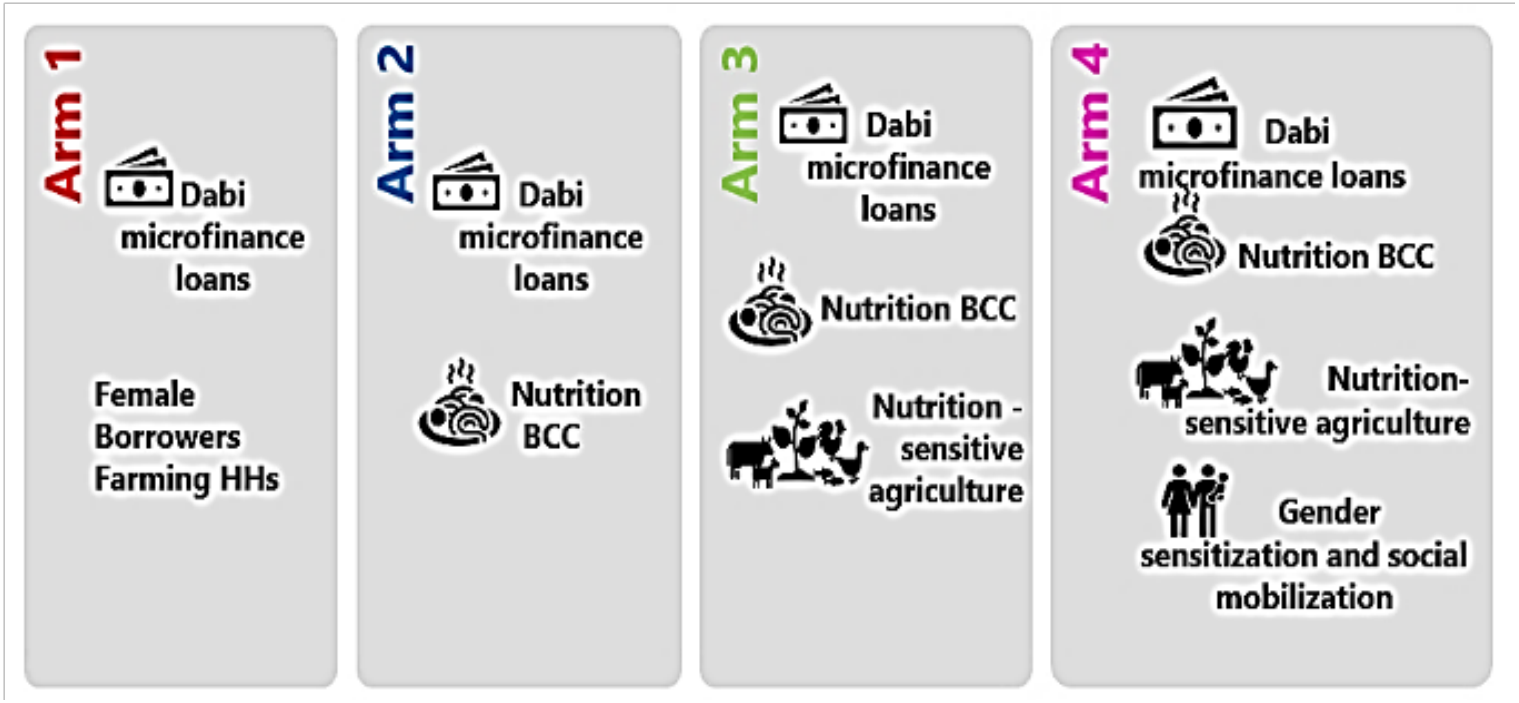


FIGURE I: TRAIN Experimental Design

Notes: The treatment group of interest is Arm 4. The control group is Arm 3.

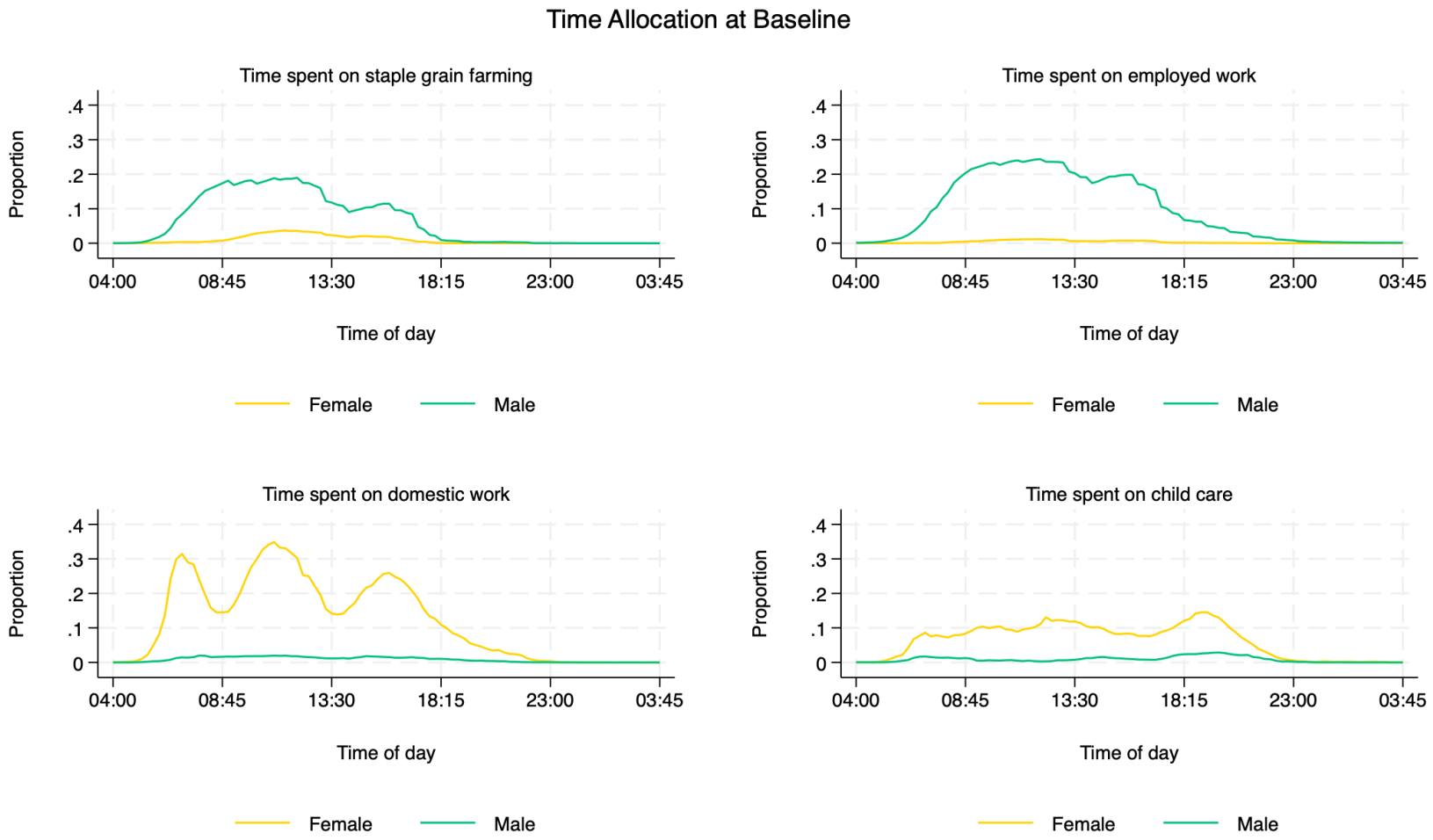


FIGURE II: Examples of Household Division of Labor

TABLE I: Power to Detect Increases in Outcomes

		T3 vs T4		
		Power to detect increase of		
Outcome	C.V.	20%	30%	40%
Share of awake time on productive work for women	1.159	0.735	0.973	0.999
Share of awake time on domestic work for men	2.660	0.178	0.340	0.541
Share of IGA decisions women involved	0.581	0.972	1.000	1.000
Share of consumption decisions women involved	0.666	0.291	0.551	0.788

Notes: IGA stands for income-generating activities. The sample is restricted to dual-adult households. C.V. stands for coefficients of variation and calculated by dividing the standard deviation by the mean. The C.V.s are derived from the endline data.

TABLE II: Summary Statistics of Baseline Household Characteristics

	Balance on the original sample			Balance on the DAHH sample		
	Treatment Arm 4 (T4)	Treatment Arm 3 (T3)	Diff p-value	Treatment Arm 4 (T4)	Treatment Arm 3 (T3)	Diff p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Household head is female	0.078	0.037	0.003	0.041	0.021	0.037
	(0.268)	(0.188)		(0.199)	(0.144)	
Age of household head	37.199	36.944	0.711	37.692	37.115	0.421
	(12.272)	(12.082)		(12.228)	(11.887)	
Years of schooling of household head	4.071	4.190	0.630	4.006	4.110	0.704
	(3.835)	(3.891)		(3.813)	(3.898)	
Household head is married	0.987	0.992	0.181	0.992	0.994	0.578
	(0.112)	(0.089)		(0.091)	(0.080)	
Household head is Muslim	0.866	0.907	0.315	0.853	0.899	0.294
	(0.341)	(0.290)		(0.354)	(0.301)	
Household head is Hindu	0.133	0.092	0.315	0.146	0.100	0.294
	(0.340)	(0.289)		(0.353)	(0.300)	
Household size	4.633	4.671	0.679	4.738	4.751	0.887
	(1.735)	(1.551)		(1.753)	(1.568)	
Ratio of male to female HH members	1.107	1.211	0.001	1.141	1.209	0.027
	(0.748)	(0.804)		(0.750)	(0.790)	
Dependency ratio	0.786	0.726	0.030	0.759	0.707	0.079
	(0.568)	(0.488)		(0.535)	(0.471)	
Number of children under 2 years old	0.287	0.304	0.348	0.284	0.296	0.512
	(0.466)	(0.472)		(0.465)	(0.469)	
Number of children under 5 years old	0.671	0.712	0.160	0.673	0.704	0.269
	(0.610)	(0.579)		(0.609)	(0.585)	

Continued on next page

Highest years of education	7.510	7.840	0.163	7.604	7.849	0.299
	(3.068)	(2.920)		(3.012)	(2.917)	
Monthly per capita non-food exp.	1,284.986	1,287.743	0.966	1,279.722	1,257.111	0.726
	(732.955)	(801.990)		(725.443)	(765.326)	
Monthly per capita HH expenditure	3,209.514	3,231.059	0.845	3,203.190	3,181.763	0.849
	(1,200.297)	(1,380.855)		(1,200.919)	(1,326.917)	
Owned homestead	0.997	0.996	0.921	0.999	0.995	0.664
	(0.164)	(0.152)		(0.165)	(0.158)	
Number of plots, incl. homestead	4.665	4.704	0.850	4.755	4.751	0.987
	(2.816)	(2.797)		(2.866)	(2.842)	
Average plot size (decimals)	26.580	24.755	0.351	26.964	25.324	0.453
	(22.844)	(19.952)		(23.581)	(20.583)	
HH food security access scale	1.189	0.979	0.254	1.124	0.988	0.439
	(2.387)	(2.397)		(2.283)	(2.447)	
Outer walls, improved materials	0.840	0.759	0.116	0.833	0.755	0.132
	(0.367)	(0.428)		(0.373)	(0.430)	
Roof, improved materials	0.966	0.983	0.152	0.964	0.982	0.177
	(0.182)	(0.128)		(0.185)	(0.135)	
Floor, improved materials	0.108	0.140	0.230	0.112	0.144	0.232
	(0.310)	(0.347)		(0.315)	(0.351)	
Dwelling has electricity	0.738	0.839	0.037	0.735	0.834	0.054
	(0.440)	(0.368)		(0.442)	(0.372)	
No. of cellular phones, working	1.663	1.721	0.348	1.690	1.739	0.452
	(1.014)	(1.003)		(1.034)	(1.012)	
Number of observations	1,260	1,260		1,067	1,083	

Notes: Columns (1) and (2) show the means for all treatment and comparison households in the study. Columns (4) and (5) show the means for treatment and comparison dual-adult households. Standard deviations are reported in parentheses. Columns (3) and (6) present the p-values from regressions of given baseline variables on a treatment indicator, indicating the differences between all treatment and comparison households and between treatment and comparison dual-adult households respectively. The regressions include strata fixed effects and cluster standard errors at the union level.

TABLE III: Summary Statistics of Baseline Characteristics of Index Women

	Balance on the original sample			Balance on the DAHH sample		
	Treatment Arm 4 (T4)	Treatment Arm 3 (T3)	Diff p-value	Treatment Arm 4 (T4)	Treatment Arm 3 (T3)	Diff p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Index woman: Age	23.717	23.343	0.032	23.636	23.311	0.093
	(3.537)	(3.572)		(3.559)	(3.572)	
Index woman: Years of schooling	6.323	6.730	0.095	6.387	6.717	0.179
	(3.203)	(2.993)		(3.117)	(2.988)	
Index woman: Married	1.000	1.000		1.000	1.000	
	(0.000)	(0.000)		(0.000)	(0.000)	
Index woman: Age at marriage	16.179	16.071	0.422	16.109	16.072	0.780
	(2.169)	(2.118)		(2.058)	(2.117)	
Index woman: Age at first birth	17.942	17.861	0.565	17.893	17.814	0.575
	(2.214)	(2.233)		(2.157)	(2.207)	
No. of children index woman has	1.408	1.363	0.320	1.411	1.361	0.328
	(0.872)	(0.836)		(0.868)	(0.850)	
Husband is co-residing	0.857	0.904	0.009	0.896	0.930	0.022
	(0.350)	(0.295)		(0.305)	(0.256)	
Mother-in-law is co-residing	0.433	0.458	0.344	0.446	0.471	0.397
	(0.496)	(0.498)		(0.497)	(0.499)	
Father-in-law is co-residing	0.313	0.336	0.391	0.333	0.355	0.417
	(0.464)	(0.472)		(0.471)	(0.479)	
Index woman is also DABI beneficiary	0.714	0.699	0.498	0.696	0.691	0.813
	(0.452)	(0.459)		(0.460)	(0.462)	
Husband's age	30.806	30.376	0.168	30.774	30.372	0.205
	(5.100)	(5.177)		(5.074)	(5.173)	
Husband's years of schooling	4.783	5.050	0.364	4.848	5.079	0.416
	(3.841)	(3.807)		(3.837)	(3.815)	
Husband: Age at marriage	22.657	22.425	0.367	22.664	22.489	0.501
	(3.375)	(3.431)		(3.370)	(3.492)	
Number of observations	1,260	1,260		1,067	1,083	

Notes: Columns (1) and (2) show the means for all treatment and comparison households in the study. Columns (4) and (5) show the means for treatment and comparison dual-adult households. Standard deviations are reported in parentheses. Columns (3) and (6) present the p-values from regressions of given baseline variables on a treatment indicator, indicating the differences between all treatment and comparison households and between treatment and comparison dual-adult households respectively. The regressions include strata fixed effects and cluster standard errors at the union level.

TABLE IV: Sample Size and Attrition Rate by Treatment Group Across Rounds

	Overall		Treatment Arm 4 (T4)		Treatment Arm 3 (T3)	
	N	Attrition	N	Attrition	N	Attrition
Baseline original household sample	5,040		1,260		1,260	
Endline original household sample	4,798	0.05	1,202	0.05	1,214	0.04
Baseline dual-adult household sample	4,783		1,171		1,221	
Endline dual-adult household sample	4,284	0.10	1,067	0.09	1,083	0.11

Notes: The attrition rate is calculated as the proportion of households in the baseline sample that were not found at endline. The original household sample includes all households enrolled at baseline. The dual-adult household sample restricts to households with both husband and wife present at baseline.

TABLE V: Selection into Non-attrition and Dual-adult Households

	Selection into Non-attrition		Selection into Endline Dual-adult Households	
	(1)	(2)	(3)	(4)
	Being a non-attrited HH	Being a non-attrited HH	Being a dual-adult HH	Being a dual-adult HH
Assigned to Gender training	0.006	0.374	-0.013	0.289
	(0.015)	(0.590)	(0.017)	(0.615)
Household size		0.035**		0.038**
		(0.013)		(0.012)
Ratio of male to female HH members		-0.016		-0.019
		(0.014)		(0.014)
Dwelling has electricity		-0.027		-0.030
		(0.026)		(0.025)
No. of cellular phones in HH		-0.014		-0.015
		(0.012)		(0.012)
Index woman is also DABI beneficiary		-0.014		-0.037
		(0.027)		(0.030)
Index woman: Age		0.039		0.052
		(0.040)		(0.039)
Index woman: Age squared		-0.001		-0.001
		(0.001)		(0.001)
Index woman: Years of schooling		0.001		-0.001
		(0.003)		(0.003)
No. of children index woman has		-0.024		-0.030
		(0.018)		(0.017)
Husband is co-residing		0.217***		0.261***
		(0.051)		(0.048)
Mother-in-law is co-residing		-0.031		-0.045
		(0.031)		(0.032)
Gender training * Household size		-0.029*		-0.024

Continued on next page

Table V – Continued from previous page

	Selection into Non-attrition		Selection into Endline Dual-adult Households	
	(1)	(2)	(3)	(4)
	Being a non-attrited HH	Being a non-attrited HH	Being a dual-adult HH	Being a dual-adult HH
		(0.014)		(0.014)
Gender training * Ratio of male to female HH members		0.042*		0.054**
		(0.018)		(0.019)
Gender training * Dwelling has electricity		0.021		0.020
		(0.036)		(0.036)
Gender training * No. of cellular phones in HH		-0.000		0.003
		(0.015)		(0.015)
Gender training * Index woman is DABI beneficiary		-0.071		-0.059
		(0.044)		(0.046)
Gender training * Index woman: Age		-0.021		-0.025
		(0.052)		(0.054)
Gender training * Index woman: Age squared		0.000		0.000
		(0.001)		(0.001)
Gender training * Index woman: Years of schooling		0.008		0.009*
		(0.004)		(0.004)
Gender training * No. of children index woman has		0.056*		0.056*
		(0.023)		(0.024)
Gender training * Husband is co-residing		-0.098		-0.002
		(0.068)		(0.066)
Gender training * Mother-in-law is co-residing		0.012		0.019
		(0.043)		(0.045)
Joint F-test RI		0.004		0.000
Control group mean	0.871	0.871	0.860	0.860
Observations	2,392	2,392	2,520	2,520
R-squared	0.013	0.046	0.016	0.096

Notes: Regressions include division dummies (not shown). Missing values replaced with means; missing indicators included. SEs clustered at union level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE VI: Summary Statistics of Program Participation

	Treatment Arm 4 (T4)	Treatment Arm 3 (T3)	Diff p-value
	(1)	(2)	(3)
Woman attended gender trainings (Mar 2019 - Mar 2020)	0.598	0.348	0.000
	(0.491)	(0.477)	
Man attended gender trainings (Mar 2019 - Mar 2020)	0.362	0.162	0.000
	(0.481)	(0.368)	
Both woman and man attended gender trainings (Mar 2019 - Mar 2020)	0.300	0.091	0.000
	(0.458)	(0.288)	
Observations	1,067	1,083	

Notes: The sample is restricted to dual-adult households found at endline. Columns (1) and (2) show the means for treatment and control households. Standard deviations are reported in parentheses. Column (3) presents the p-values from regressions of given variables on a treatment indicator, indicating the differences between treatment and control households. The regressions include strata fixed effects and cluster standard errors at the union level.

TABLE VII: Impact of Gender Trainings on Exposure to Gender Messages

	(1) Woman discussed division of labor	(2) Man discussed division of labor	(3) No. of gender topics woman recognized	(4) No. of gender topics woman discussed	(5) No. of gender topics man recognized	(6) No. of gender topics man discussed
Assigned to T2	0.059**	0.028*	1.089***	0.653***	0.496**	0.266**
	(0.028)	(0.015)	(0.297)	(0.241)	(0.234)	(0.126)
Assigned to T3	0.201***	0.109***	2.063***	1.657***	1.277***	0.508***
	(0.034)	(0.022)	(0.243)	(0.250)	(0.262)	(0.145)
Assigned to T4	0.393***	0.233***	2.864***	2.578***	1.947***	1.085***
	(0.041)	(0.029)	(0.255)	(0.227)	(0.242)	(0.151)
R^2	0.140	0.091	0.166	0.161	0.110	0.074
P-value T3 = T1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0006
P-value T3 = T2	0.0005	0.0011	0.0008	0.0006	0.0035	0.1376
P-value T3 = T4	0.0002	0.0005	0.0010	0.0009	0.0142	0.0018
Control group mean	0.067	0.020	1.879	0.950	1.390	0.448
Observations	4,284	4,284	4,284	4,284	4,284	4,284
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes
PDS Lasso-selected controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The sample is restricted to dual adult households at endline. Control variables that feed into post-double Lasso include the corresponding baseline outcome, household characteristics in Tables 2 and 3. When a household has missing in a baseline control variable, the missing value is replaced with the mean value for the sample, and a dummy variable is included for all households for which the baseline value is missing. The coefficients of the selected control variables are not shown here. Regressions also include division dummy variables and are adjusted for weights that combine selective attrition weights and selection into dual-adult-household weights. Standard errors are clustered at the union level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE VIII: Impact of Gender Trainings on Household Division of Labor

	Time Use		Decision Making	
	(1) Share of awake time on productive work for women	(2) Share of awake time on domestic work for men	(3) Share of productive decisions women involved	(4) Share of consumption decisions women involved
Both spouses participated in gender training	-0.007	0.031	-0.079	-0.025
	(0.040)	(0.026)	(0.104)	(0.063)
Effective F stats	40.093	38.301	38.756	38.553
Control group mean	0.118	0.026	0.419	0.258
Observations	2,150	2,150	2,150	2,150
Strata FE	Yes	Yes	Yes	Yes
PDS Lasso-selected controls	Yes	Yes	Yes	Yes

Notes: The sample is restricted to dual adult households at endline. Control variables that feed into post-double Lasso include the corresponding baseline outcome, household characteristics in Tables 2 and 3. When a household has missing in a baseline control variable, the missing value is replaced with the mean value for the sample, and a dummy variable is included for all households for which the baseline value is missing. The coefficients of the selected control variables are not shown here. Regressions also include division dummy variables and are adjusted for weights that combine selective attrition weights and selection into dual-adult-household weights. Standard errors are clustered at the union level. Stars denote statistical significance as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE IX: Model Performance Metrics

	Share of awake time on productive work for women		Share of awake time on domestic work for men		Share of productive decisions women involved		Share of consumption decisions women involved	
Learner	λ	$\bar{\lambda}$	λ	$\bar{\lambda}$	λ	$\bar{\lambda}$	λ	$\bar{\lambda}$
Random forest	0.00071	0.00111	0.00020	0.00026	0.00989	0.00876	0.00311	0.00284
Support vector machine	0.00095	0.00125	0.00001	0.00010	0.01955	0.01507	0.00541	0.00405
Neural net	0.00021	0.00052	0.00003	0.00016	0.00983	0.00924	0.00143	0.00180
Gradient boosting	0.00060	0.00101	0.00022	0.00030	0.00523	0.00555	0.00259	0.00264

TABLE X: Model Estimates and Confidence Intervals

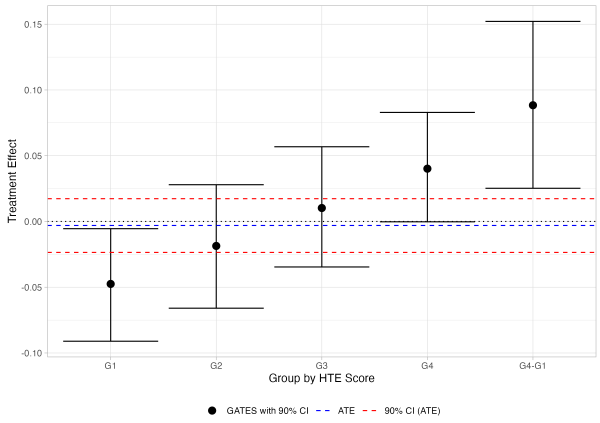
	Share of awake time on productive work for women				Share of awake time on domestic work for men				Share of productive decisions women involved				Share of consumption decisions women involved			
Parameter	Estimate	CI Lower	CI Upper	P-Value	Estimate	CI Lower	CI Upper	P-Value	Estimate	CI Lower	CI Upper	P-Value	Estimate	CI Lower	CI Upper	P-Value
beta.1	-0.0030	-0.0235	0.0173	0.7731	0.0057	0.0053	0.0166	0.3106	-0.0072	-0.0418	0.0270	0.6720	-0.0026	-0.0298	0.0221	0.8393
beta.2	1.2814	0.3350	2.2259	0.0081	0.5167	0.0819	0.9546	0.0193	1.3112	0.8999	1.7285	0.0000	1.0378	0.6082	1.4702	0.0000

Notes: The table presents the estimated coefficients, 95% confidence intervals (CI Lower/CI Upper), and p-values for two parameters (β_1 and β_2) across four different outcome variables related to time use and decision-making in households. β_1 indicates average treatment effects while β_2 test the null hypothesis of no heterogeneity.

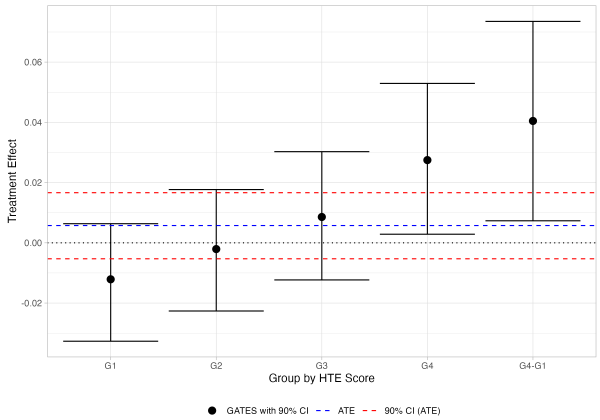
TABLE XI: Heterogeneous Treatment Effects by Quartile

	Share of awake time on productive work for women	Share of awake time on domestic work for men	Share of productive decisions women involved	Share of consumption decisions women involved
Estimates and 95% Confidence Intervals				
First quartile	-0.048^{**}	-0.012	-0.211^{***}	-0.088^{***}
	$[-0.091, -0.006]$	$[-0.033, 0.006]$	$[-0.310, -0.112]$	$[-0.145, -0.032]$
Second quartile	-0.020	-0.002	0.054	-0.028
	$[-0.066, 0.028]$	$[-0.023, 0.018]$	$[-0.059, 0.071]$	$[-0.076, 0.021]$
Third quartile	0.010	0.009	0.050	0.019
	$[-0.035, 0.057]$	$[-0.012, 0.030]$	$[-0.016, 0.113]$	$[-0.033, 0.071]$
Fourth quartile	0.040^{*}	0.028^{**}	0.116^{***}	0.080^{***}
	$[-0.000, 0.083]$	$[0.003, 0.053]$	$[0.039, 0.194]$	$[0.024, 0.136]$
Difference (Q4 - Q1)	0.089^{***}	0.041^{**}	0.324^{***}	0.168^{***}
	$[0.025, 0.152]$	$[0.007, 0.074]$	$[0.184, 0.467]$	$[0.080, 0.256]$

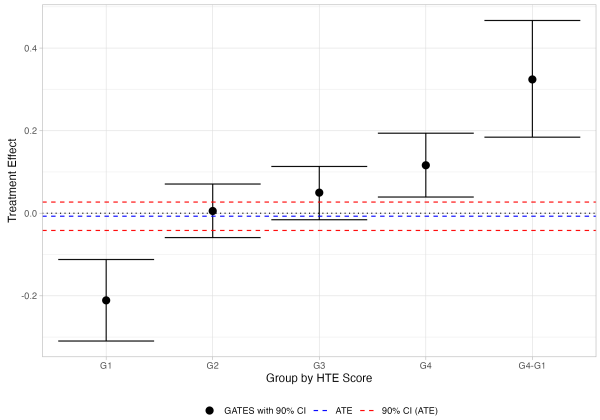
Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



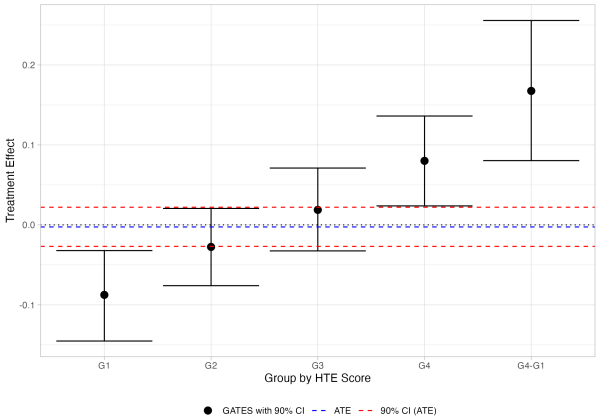
(A) Panel A: Share of women's time on productive work



(B) Panel B: Share of men's time on domestic work



(C) Panel C: Share of IGA decisions women involved



(D) Panel D: Share of consumption decisions women involved

FIGURE III: Heterogeneous Treatment Effects

Appendix Table I: Description of Outcomes

Variable	Description
Time use	
Share of awake time on paid work for women	Total minutes woman spent on paid work at endline divided by total minutes woman stayed awake (1440 minutes minus minutes on sleeping). Activities under paid work include salaried work, own business work, staple grain farming, high value crop, large livestock raising (cattle, buffaloes), small livestock raising (sheep, goats, pigs), poultry and other small animals raising (chickens, ducks, turkeys), and fishpond culture.
Share of awake time on domestic work for women	Total minutes woman spent on domestic work at endline divided by total minutes woman stayed awake (1440 minutes minus minutes on sleeping). Activities under domestic work include weaving / sewing / textile care, cooking, domestic chores (fetching water and collecting fuel), caring for children, and caring for adults.
Share of awake time on paid work for men	Total minutes man spent on paid work at endline divided by total minutes man stayed awake (1440 minutes minus minutes on sleeping). Activities under paid work include salaried work, own business work, staple grain farming, high value crop, large livestock raising (cattle, buffaloes), small livestock raising (sheep, goats, pigs), poultry and other small animals raising (chickens, ducks, turkeys), and fishpond culture.
Share of awake time on domestic work for men	Total minutes man spent on domestic work at endline divided by total minutes man stayed awake (1440 minutes minus minutes on sleeping). Activities under domestic work include weaving / sewing / textile care, cooking, domestic chores (fetching water and collecting fuel), caring for children, and caring for adults.
Decision Making	
Share of IGA decisions women involved	Total number of IGA decisions women involved at endline divided by the total number of IGA household participated in. IGA include staple grain farming, high value farming, large livestock raising, small livestock raising, poultry and other small animals raising, fishpond culture, non-farm economic activities, and wage and salary employments.
Share of consumption decisions women involved	Total number of decisions related to household consumption women involved at endline divided by the total number of consumption decisions. They include large occasional household purchases and routine household purchases.
Share of IGA decisions men involved	Total number of IGA decisions men involved at endline divided by the total number of IGA household participated in. IGA include staple grain farming, high value farming, large livestock raising, small livestock raising, poultry and other small animals raising, fishpond culture, non-farm economic activities, and wage and salary employments.
Share of consumption decisions men involved	Total number of decisions related to household consumption men involved at endline divided by the total number of consumption decisions. They include large occasional household purchases and routine household purchases.

Appendix Table II: Description of baseline household characteristics

Variable	Description
Dependency ratio	Percentage of individuals under 15 and over 64 years old within the household
Household food insecurity access scale score	Sum of the frequency-of-occurrence during the past four weeks for the 9 food insecurity-related conditions, range from 0 to 27. The higher the score, the more food insecurity the household experienced (USAID, 2007)
Outer walls of dwelling made with improved materials	Value of 1 if outer walls of dwelling made with concrete or brick or tin or CI sheet, and 0 otherwise
Roof of dwelling made with improved materials	Value of 1 if floor of dwelling made with concrete or brick or tin or CI sheet, and 0 otherwise
Floor of dwelling made with improved materials	Value of 1 if floor of dwelling made with concrete or brick, and 0 otherwise

Appendix Table III: Baseline Values for Bottom and Top Quartiles

	Share of awake time on productive work for women			Share of awake time on domestic work for men			Share of productive decisions women involved			Share of consumption decisions women involved		
Variable	δ_1	δ_{diff}	p-value	δ_1	δ_{diff}	p-value	δ_1	δ_{diff}	p-value	δ_1	δ_{diff}	p-value
Female employment rate	0.0307	-0.0012	0.5571	0.0249	0.0159	0.0000	0.0262	0.0091	0.0000	0.0247	0.0128	0.0000
Woman had some input in productive decision in staple grain farming	24.0344	1.4793	0.0031	23.9751	0.8072	0.1157	23.2572	1.8863	0.0001	24.7514	-0.5658	0.2517
Woman had some input in productive decision in high value crop farming	6.8969	-1.5164	0.0012	5.9962	0.2806	0.5251	3.7364	2.6051	0.0000	4.6419	0.7678	0.0497
Woman had some input in productive decision in large livestock	16.8686	-0.8975	0.0663	15.2495	1.1514	0.0155	15.6164	0.8151	0.0521	15.5625	0.0960	0.8302
Woman had some input in productive decision in small livestock	7.8037	0.6381	0.1230	9.3813	-1.7256	0.0001	6.8423	2.5064	0.0000	7.4811	1.8815	0.0001
Woman had some input in productive decision in poultry and other small animals raising	15.5432	1.1633	0.0078	14.4589	2.0418	0.0000	12.7670	3.6452	0.0000	14.3066	1.8325	0.0001
Woman had some input in productive decision in fishpond culture	1.8211	0.0598	0.8607	2.0193	-0.3110	0.2767	1.4453	0.8837	0.0108	1.6053	0.4289	0.1335
Woman had some input in productive decision in non-farm activities	2.4432	0.4983	0.1785	2.2844	1.2575	0.0007	3.2004	-0.9947	0.0026	4.3075	-2.4034	0.0000
Woman had some input in productive decision in wage and salary employment	1.8356	1.0656	0.0003	2.0350	0.8060	0.0055	1.6483	1.1415	0.0000	2.7703	-0.8576	0.0063
Number of productive decisions woman had some inputs	11.8210	-0.5356	0.1553	11.0158	0.2645	0.4677	10.0058	2.0124	0.0000	10.1984	0.5457	0.1292
Woman solely or jointly owned large livestock	7.3897	1.3080	0.0011	9.6800	-2.0561	0.0000	6.4377	3.5580	0.0000	6.6149	2.4641	0.0000
Woman solely or jointly owned small livestock	24.3562	0.6601	0.0173	24.8516	-0.5534	0.0500	23.7381	0.7002	0.0297	24.5145	-0.1963	0.5114
Woman solely or jointly owned poultry and other small animals	3.5603	-0.3285	0.3739	4.6309	-1.3594	0.0034	3.6641	-0.1577	0.7365	2.5461	0.6572	0.0518
Woman solely or jointly owned fishpond or fishing equipment	14.9168	-1.3941	0.0775	14.6238	-0.8085	0.3254	14.0605	-0.5930	0.4511	13.9216	-0.7374	0.3695
Woman solely or jointly owned non-mechanized farm equipment	1.7084	-0.6321	0.0000	1.5413	0.0159	0.9130	1.4577	0.5735	0.0001	1.3612	0.1518	0.2635
Woman solely or jointly owned mechanized farm equipment	4.1542	-0.6017	0.0574	4.8008	-0.8950	0.0256	4.4344	-1.1219	0.0019	3.4068	-0.0702	0.7793
Woman solely or jointly owned non-farm business equipment	7.7298	-0.9265	0.1233	7.6773	-0.9294	0.1403	5.7415	2.5227	0.0000	6.5759	1.7268	0.0033
Woman solely or jointly owned housing or building	10.5768	1.2769	0.0052	10.3465	1.9902	0.0000	9.1811	2.9461	0.0000	10.1663	1.8279	0.0002
Woman solely or jointly owned large consumer durables	29.9656	-1.1772	0.0036	30.2731	-1.3512	0.0008	30.5105	-2.3989	0.0000	29.8075	-1.7620	0.0000
Woman solely or jointly owned small consumer durables	21.0207	0.5082	0.2667	21.5184	-0.0862	0.8591	20.2694	1.4554	0.0006	21.1279	0.7987	0.1012
Woman solely or jointly owned a cell phone	0.3829	-0.0765	0.3317	0.4215	0.0035	0.9733	0.4674	-0.1832	0.0184	0.7231	-0.5267	0.0000
Woman solely or jointly owned other land not used for agricultural purposes	5.1055	-0.6951	0.0720	5.3887	-0.5952	0.1589	4.9374	0.8696	0.0410	4.6269	0.7910	0.0721
Woman solely or jointly owned means of transportation	2.7816	0.6961	0.0252	2.9679	0.0745	0.8151	2.1823	1.6133	0.0000	3.1537	0.6877	0.0386
Woman solely or jointly owned land	28.0094	-1.3177	0.0123	27.7562	-0.3756	0.4211	27.8872	-0.7426	0.0810	27.7649	0.0780	0.8649
Number of assets women solely or jointly owned	2.7937	-0.6424	0.0106	2.5764	0.2354	0.3858	3.3558	-1.0879	0.0000	3.1122	-0.3558	0.1824
Woman used or made decision in credit from NGO	2.2245	0.3252	0.2444	2.4212	0.4969	0.0994	1.4079	1.4793	0.0000	1.8547	1.5106	0.0000
Woman used or made decision in credit from informal lender	9.8943	-1.1526	0.0279	10.3296	-1.1229	0.0424	11.0931	-3.6413	0.0000	10.4730	-2.1798	0.0000
Woman used or made decision in credit from formal lender	2.3371	-0.6184	0.0462	2.8528	-0.4663	0.2252	2.4038	-0.3079	0.3517	1.2916	1.0096	0.0003
Woman used or made decision in credit from friends or relatives	0.3755	-0.0729	0.2415	0.2731	0.0833	0.1413	0.3707	-0.0004	0.9947	0.3931	-0.0652	0.3646
Woman used or made decision in credit from group-based microfinance	10.2320	0.9502	0.1348	10.5560	-0.7763	0.1950	8.4004	3.1773	0.0000	8.9316	2.7391	0.0000
Woman used or made decision in credit from informal savings groups	23.4668	1.7253	0.0006	23.4934	1.2092	0.0153	22.9791	1.8060	0.0002	24.2313	-0.4887	0.3309
Woman had a bank account	6.7266	-1.1986	0.0097	5.6521	0.7864	0.0770	3.6565	2.6642	0.0000	4.5880	0.4155	0.2494
Number of credits women had access to	16.3809	-0.3906	0.3416	14.8707	1.7590	0.0001	15.5808	0.9515	0.0114	15.1756	0.2871	0.4767
Woman had some input in income decision in staple grain farming	8.2519	0.8315	0.0369	10.0266	-1.7029	0.0000	7.5511	2.5323	0.0000	7.7170	2.2188	0.0000
Woman had some input in income decision in high value crop farming	25.3233	0.8811	0.0033	25.4477	0.4989	0.1142	25.1518	0.5578	0.0704	25.4677	0.1450	0.6606
Woman had some input in income decision in large livestock	1.7439	0.0789	0.7673	1.9251	-0.2733	0.3364	1.3080	0.9510	0.0052	1.5056	0.3758	0.1616
Woman had some input in income decision in small livestock	2.5992	0.6366	0.0995	2.4827	1.4098	0.0003	3.6391	-1.3329	0.0001	4.6502	-2.6198	0.0000
Woman had some input in income decision in poultry and other small animals raising	2.0702	1.0954	0.0003	2.3477	0.7297	0.0157	1.7718	1.2789	0.0000	2.8825	-0.7089	0.0301
Woman had some input in income decision in fishpond culture	67.6485	-7.9727	0.0000	66.9938	-6.7809	0.0001	69.0425	-4.2746	0.0119	64.5190	0.2181	0.9082
Woman had some input in income decision in non-farm activities	4.2355	-1.6954	0.0009	4.9036	-1.3196	0.0339	6.2852	-3.3430	0.0000	3.5904	-1.2352	0.0026
Woman had some input in income decision in wage and salary employment	1.2645	0.0159	0.9305	1.3078	0.3369	0.2001	1.7681	-0.3879	0.1328	0.8203	0.5855	0.0073
Number of income decisions woman had some inputs	0.1523	0.0000	1.0000	0.3801	-0.3280	0.0000	0.4192	-0.3687	0.0000	0.0920	-0.0536	0.0187
Woman's autonomy in income score	18.0136	-1.8185	0.0157	18.0838	-1.5561	0.0329	16.7669	-0.3528	0.6218	16.5626	-1.9650	0.0092
Woman was a member of agricultural producer's group	2.5895	-1.3720	0.0000	2.5569	-0.5013	0.1364	3.3643	-2.1268	0.0000	2.4310	-1.2915	0.0000
Woman was a member of water users' group	5.3117	-2.5570	0.0000	5.5764	-1.6466	0.0076	7.5451	-4.8501	0.0000	5.3985	-2.9317	0.0000
Woman was a member of forest users' group	0.9886	-0.5302	0.0018	1.2095	-0.0457	0.8441	1.2404	-0.8122	0.0000	0.6107	-0.1552	0.1616

Continued on next page

Table XII – continued from previous page

	Share of awake time on productive work for women			Share of awake time on domestic work for men			Share of productive decisions women involved			Share of consumption decisions women involved		
Variable	δ_1	δ_{diff}	p-value	δ_1	δ_{diff}	p-value	δ_1	δ_{diff}	p-value	δ_1	δ_{diff}	p-value
Woman was a member of microfinance group	10.4635	-1.3930	0.0477	7.9181	3.8634	0.0000	11.7121	-2.7419	0.0001	12.6087	-3.0799	0.0000
Woman was a member of insurance group	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
Woman was a member of trade and business association group	0.0410	-0.0380	0.0013	0.0542	-0.0526	0.0001	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
Woman was a member of civic group	0.0686	-0.0533	0.0001	0.0587	0.0343	0.2065	0.1853	-0.1841	0.0000	0.0000	0.0000	1.0000
Woman was a member of religious group	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
Woman was a member of other group	5.1099	-1.8065	0.0000	4.6030	-0.1451	0.7236	4.3037	0.0046	0.9907	3.5627	0.0092	0.9802
Number of groups woman was a member	0.0587	-0.0576	0.0001	0.0000	0.0525	0.0001	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
Woman was an influential member of agricultural producer's group	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
Woman was an influential member of water users' group	0.0414	-0.0342	0.0023	0.0000	0.0971	0.0000	0.0000	0.0359	0.0013	0.0000	0.0000	1.0000
Woman was an influential member of forest users' group	0.8650	-0.4393	0.0001	0.4886	0.6761	0.0007	0.5993	0.4434	0.0178	0.4713	0.1781	0.0530
Woman was an influential member of microfinance group	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000
Woman was an influential member of insurance group	16.6262	1.4276	0.0015	17.5744	0.1780	0.6692	14.8480	2.5618	0.0000	15.7273	2.2319	0.0000
Woman was an influential member of trade and business association group	0.2413	0.3563	0.0000	0.2381	0.4782	0.0000	0.2519	0.3044	0.0000	0.3219	0.3173	0.0000
Woman was an influential member of civic group	1.1219	-0.2600	0.0378	0.8067	0.5779	0.0000	1.1399	-0.1817	0.1765	1.1042	0.0307	0.8498
Woman was an influential member of religious group	3.3957	-0.9665	0.0002	2.8095	0.1174	0.6571	3.8805	-1.0869	0.0001	3.7887	-0.6109	0.0445
Woman was an influential member of other group	6.2892	-0.9827	0.0274	6.8561	-1.2597	0.0058	6.1635	0.8264	0.0783	6.2773	0.4699	0.3397
Number of groups woman was an influential member	6.0719	-0.9507	0.0237	5.5592	0.9438	0.0266	7.5524	-1.7026	0.0001	6.5695	-1.2811	0.0028
Woman had balanced workload	33.4416	-0.9643	0.0007	33.4308	-1.1538	0.0001	32.2404	-0.0578	0.8815	33.1279	-0.4466	0.1462
Hours of work woman had	30.2476	0.3565	0.3586	30.6004	-0.6268	0.1008	29.5665	0.6737	0.1119	30.6223	-0.7505	0.0587
Woman visited weekly an urban center	29.0821	-0.5353	0.2251	29.6329	-2.4315	0.0000	28.8280	-1.2133	0.0087	29.6211	-2.6958	0.0000
Woman visited weekly a market	21.2955	0.0194	0.9734	21.4137	-1.4578	0.0111	20.3526	-0.4869	0.3756	21.1924	-2.0239	0.0002
Woman visited weekly family	29.1560	-0.8991	0.0140	29.9075	-1.9331	0.0000	29.6509	-0.2378	0.4992	29.0716	0.7764	0.0298
Woman visited monthly the hospital	29.6904	-2.0163	0.0000	29.9981	-2.1819	0.0000	29.4099	-0.0047	0.9888	28.7917	0.4443	0.2106
Woman visited monthly public gathering	25.5981	-0.4892	0.2488	25.9534	-0.4841	0.2514	26.0945	0.4166	0.2364	26.0792	0.5355	0.1565
Number of places woman visited	33.2879	-0.8504	0.0002	33.0863	-0.7302	0.0019	33.0112	0.1123	0.5967	32.9253	-0.1098	0.6047
Woman respected her family most of the time	32.8696	0.1186	0.5094	32.9739	-0.0414	0.8448	33.1868	0.3737	0.0248	33.2467	-0.0853	0.6516
Woman was respected by her family most of the time	129.8175	0.0227	0.9803	129.7500	-0.0027	0.9980	128.1465	6.0050	0.0000	128.5412	3.4272	0.0000
Woman trusted her family most of the time	132.7871	-1.3522	0.0883	131.2198	2.1168	0.0110	131.4031	3.8794	0.0000	131.5433	1.3597	0.0916
Woman felt comfortable telling her family that she disagreed most of the time	134.4779	-0.7017	0.3453	133.4431	0.9268	0.2449	133.9043	3.0059	0.0000	133.9319	0.6833	0.3702
Woman's respect among household members	129.9367	-0.4526	0.6490	128.2713	0.6440	0.5189	128.4677	4.5329	0.0000	130.1518	0.5002	0.5682
Woman: it is not justified that a husband hit or beat his wife if she goes out without telling him	77.4855	2.3199	0.1645	75.5996	4.2762	0.0089	68.6464	12.5047	0.0000	75.3762	1.0828	0.4888
Woman: it is not justified that a husband hit or beat his wife if she neglects the children	145.2245	-2.0919	0.5115	149.7121	-6.3116	0.0544	137.0934	11.8396	0.0003	139.0151	6.1691	0.0504
Woman: it is not justified that a husband hit or beat his wife if she argues with him	55.8521	-2.3761	0.0558	56.8115	-1.9663	0.1116	54.9828	-1.1507	0.3211	53.8543	2.8105	0.0106
Woman: it is not justified that a husband hit or beat his wife if she refuses to have sex with him	86.8610	3.5000	0.0130	86.4318	4.3502	0.0022	81.7552	9.3183	0.0000	86.2841	-0.2453	0.8517
Woman: it is not justified that a husband hit or beat his wife if she burns the food	11.8427	-2.4832	0.0001	10.8479	0.4787	0.4913	10.6343	-0.5548	0.4219	10.9238	-1.2510	0.0496
Number of scenarios that it is not justified for violence	6.2247	-2.3928	0.0000	5.2390	0.6027	0.2909	5.1250	0.2868	0.5595	4.0750	0.2938	0.4776
When facing difficult tasks, I am certain that I will accomplish them	368.9631	-10.7563	0.0006	363.0261	-2.0447	0.5211	383.9079	-22.9069	0.0000	376.9784	-20.8574	0.0000
In general, I think that I can obtain outcomes that are important to me	17.1085	-2.7503	0.0019	16.3118	0.8967	0.3235	19.0388	-1.8658	0.0599	18.0897	-1.0977	0.2872
I am confident that I can perform effectively on many different tasks.	114.1120	-1.0819	0.4125	115.0000	-5.5581	0.0001	110.8707	-1.1820	0.4578	114.6063	-5.9145	0.0000
Even when things are tough, I can perform quite well.	150.6128	-4.1591	0.0010	151.9110	-5.2082	0.0000	151.2106	0.6442	0.5739	150.0687	1.6039	0.1930
Number of self-efficacy statements	526.9750	-2.5207	0.4569	522.4280	3.6147	0.2769	521.9589	17.3598	0.0000	524.2867	6.0493	0.0450
Woman was empowered	6.4629	-0.7538	0.0398	6.1316	0.0898	0.8168	5.1445	0.9847	0.0036	5.2524	0.6187	0.0558

Notes: δ_1 indicates the baseline value of the bottom quartile while δ_{diff} refers to the difference in the baseline value between the bottom and the top quartile.