AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Investment and Household Bargaining in Small-scale Farming Households A Lab-in-the-Field Experiment from Rural Tanzania

by Michael Keenan, Haki Pamuk, Stanley Karanja Ng'ang'a, and Joseph Phillip Hella

Copyright 2021 by Michael Keenan, Haki Pamuk, Stanley Karanja Ng'ang'a, and Joseph Phillip Hella. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Investment and Household Bargaining in Small-scale Farming Households
A Lab-in-the-Field Experiment from Rural Tanzania

Michael Keenan[1,2], Haki Pamuk[1,3,*], Stanley Karanja Ng'ang'a[2], Joseph Phillip Hella[4]

[1]: Development Economics Group, Wageningen University and Research,
[2]: The Alliance of Bioversity International and CIAT
[3]: Wageningen Economic Research, Wageningen University and Research,
[4]: Department of Food and Natural Resources, Sokoine University of Agriculture.
* Corresponding author, haki.pamuk@wur.nl

Abstract: Agricultural investments often carry labor costs and market risks, and spouses negotiate and bargain to settle their conflicting preferences for risk and differing views on individual labor costs. Using a lab-in-the-field experiment with a real-effort task, we analyze man, woman, and joint decision-making in those investments. We find men prefer to take production and market risks more often than women, investing higher amounts in risky investments. However, those investments' returns are low, as they cannot complete the task in which they have invested. When couples jointly decide, they choose investment amounts in line with man spouse preferences; however, allocate investments between risky and not risky investments in line with woman preferences. We also find that woman spouses who have participated in a gender-related training program take more risks and the joint investment decisions of the couple are closer to woman investment decisions when compared to other couples. Our results are in line with previous studies on the behavioral aspects of agricultural investments and can guide organizations in improving program design to increase agricultural technology and crop adoption. This paper contributes to the scarce literature on behavioral aspects influencing the gender gap in agricultural production and technology adoption.

Keywords: Gender, Household Bargaining, Agricultural Investments, Gender Training, savings and loan associations.

Acknowledgements: We thank Prof Dr. Ruerd Ruben and Marcel van Asseldonk from Wageningen University and Research, Edgar Begasha, Blandina Karoma, and Thabit Masoud from CARE-International in Tanzania, Karl Deering from CARE-USA, and Merlijn van Waas from CARE-Netherlands for their comments and contributions to the study. This working is part of the work conducted under the CSA/SuPER - Upscaling CSA with small-scale food producers organised via VSLAs under the CCAFS program project. The authors acknowledge both financial and technical support from 4th Global Challenges Programme (GCP 4) – a joint climate smart agriculture scaling program implemented in Eastern and Southern AFRICA with the support from the Netherlands Organization for Scientific Research (NWO) and the CGIAR Research Program on Climate Change, Agriculture and

Food Security (CCAFS) for East Africa, which is carried out with support from the CGIAR Trust Fund and through bilateral funding agreements. For details please visit https://ccafs.cgiar.org/donors. The research also benefited from the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), with support from CGIAR Fund Donors and through bilateral funding agreements. This study also benefits from the CGIAR Research Program on Policies Institutions and Markets which has provided funding for Michael Keenan's PhD dissertation, which includes this working paper.

1. Introduction

The majority of the workforce in Sub-Saharan Africa (SSA) is employed in the agricultural sector (International Labour Organization, 2021), and small-scale farming accounts for up to 80% of agricultural production in SSA (IFAD, 2021). Despite the importance of small-scale farming, growth in small-scale productivity levels has remained stagnant in the past decades (Udry, 2010). While there are many factors determining farm production levels (environment, institutions, climate, technology, etc.), one important crosscutting issue is gender. Women account for about 40% of labor across SSA with considerable heterogeneity across countries. For example, in Tanzania women account for 52% of agricultural labor (Palacios-Lopez et al., 2017). Yet, there are large productivity differences between men and women across SSA. Using ISA-LSMS data, UNDP (2015) finds that the conditional gender gap in productivity is around 30% in Uganda, Malawi, and Tanzania.

The gender gap in agricultural productivity is impeding progress towards reaching the Sustainable Development Goals (SDGs). Across Uganda, Malawi, and Tanzania, closing the gender gap would result in nearly 500,000 people escaping poverty and nearly 300 million USD added to annual GDP (UNDP, 2015). Nutritional outcomes are also compromised as a result. Empowerment of women leads to better nutritional outcomes for children because women are more likely to spend money on shared family resources compared to private goods (Meinzen-dick et al., 2014). Further, increasing agricultural productivity is associated with better nutritional outcomes (Pauw and Thurlow, 2011). Closing the gender gap would also lead to greater women's empowerment and lead to a more gender equal SSA (van den Bold et al., 2013).

The gender gap in agricultural productivity can be attributed to a range of factors. Typically across SSA, women have less access to inputs (Gilbert et al. 2002, Meinzen-Dick et al. 2014), transportation (Ali et al., 2016), finance (Adegbite and Machethe, 2020), and agricultural extension (Gilbert et al. 2002, Manfre et al. 2013); lower fulfillment of land rights (Meinzen-dick et al. 2014, Yokying and Lambrecht 2020); larger time burdens in terms of agricultural labor, unpaid household duties, and child care (Palacios-Lopez and Lopez, 2014); less household bargaining power (Huyer, 2016); lower rates of technology adoption (Nyasimi and Huyer, 2017) and cash crop adoption (Ali et al., 2016); and higher vulnerability to climate change (Goh 2012, Van Aelst and Holvoet 2016). In recent decades, governments and international organizations have been trying to bridge the gender gap through gender-sensitive agricultural programs that address some of the issues listed above (FAO, 2013).

Agricultural programs are designed to improve yields and build resilience to climate change impacts, often through promoting investments in new practices/approaches. These practices are usually new to the farmers, involve production and market risk, and require new household labor allocation. If their adoption improves among women, the gender gap can be reduced, potentially reducing the overall poverty level. For example, in the 2000s, the Malawian government ran the Farm Input Subsidy Program targeting half of its farmers. Households received vouchers for improved seeds and inorganic fertilizers. Fisher and Kandiwa (2014) looks at the differential effects of the program on woman household heads (typically unmarried, divorced, widowed, or the husband is away for six months or more each year) and married women living with man household heads. The program caused a 222% increase in improved maize adoption in plots managed by woman household heads. However, plots managed by married women living with man household heads did not increase improved maize adoption. While this experience can guide future policy on closing the gender gap for woman household heads, it was ineffective in improving married women's adoption of technology. Because most woman-headed plots are managed by married women, the program appeared to have little effect on closing the gender productivity gap overall (Karamba and Winters, 2015). However, married women make up a large portion of women in SSA, living households as 'man-headed' (Buvinic and Gupta, 1997). So gender-sensitive programs need to effectively target married women as well as woman household heads, addressing

intrahousehold power dynamics and differences between men's and women's preferences (Doss, 2001; Doss et al., 2018; Rola & Rubzen et al., 2020).

In the case of married couples, particularly, the differences between men's and women's risk preferences and assessment of their own and their spouse's labor costs might influence the investment in new agricultural practices. For instance, it has been shown that women, including women in agriculture, tend to be more risk-averse than men (Vieider et al., 2015), and are thus less likely to make risky investments. Women in developing countries also have more domestic duties and perform more labor-intensive agricultural tasks, such as sowing, weeding, and harvesting (Heise et al., 2019). This could increase their marginal cost of labor (in terms of time). Since patriarchal norms call for women to provide more labor in household activities than men, men potentially undervalue their spouse's time while women overvalue their spouse's time.

In light of these divergent behavioral preferences affecting agricultural investment, who makes investment decisions in small-scale agriculture (e.g. men, women, or both), and who works as a result of investment within the household might influence investment levels and outcomes of agricultural practices. The negotiating and bargaining process between men and women determines household investment strategies when there are conflicting preferences for potential investments. Understanding this bargaining process and the effects of women's involvement in decision making is crucial to designing programs to bridge the gender gap particularly.

To understand the bargaining process and its effect of the women's involvement in decision growing number of the research study the role of women in agricultural households decision and livelihood. For instance, Sheahan & Barrett, 2014 use data from six Sub-Saharan Africa and show that woman headed are less likely to adopt agricultural input than man headed households. Shibata et al. (2020) utilized data from rural farmers in Uganda and find that adoption of agricultural innovations involve bargaining between spouses, and the power of the spouses in this bargaining is related to the social norms and their control over household assets. Ngoma-Kasanda and Sichilima (2016) provide similar evidence from groundnut farmers in Zambia, reporting that household assets, distance to the field, household headship,

cultural factors influence women's decision making in the household. Kawarazuka et al. (2018) use information from Ede community in Vietnam and find that men farmers have a higher power in farming decisions than their wives because they control farming machines (two-wheel tractors), and men farmers are not aware of the investment needs of women. There is also evidence suggesting that the involvement of women in household decision making is positively related to nutrition, livelihood, wellbeing, and resilience outcomes (please see Njuki et al. (2021) for a detailed review). A recent study by the CARE impact evaluation team uses the randomized controlled trial to test whether a gender transformative approach empowering women in the household decision can improve adoption of agricultural technologies, household dietary diversity, and income in Burundi. The study found that empowering women improved dietary diversity and contributed to the improvement of other outcomes (CARE, 2021).

Our study contributes to this growing literature, studying how women's involvement and household bargaining change decisions on risky agricultural investments requiring labor inputs of spouses. Specifically, we use a lab-in-the-field experiment with a real-effort task to measure rural man and woman spouses' different investment outcomes and compare them with their joint decision when labor costs and risks are present. Our sample includes 539 married couples across fifteen villages in Iringa, Tanzania. Rural Tanzania and particularly the villages that we focus on is a suitable setting to analyze spousal investment decision-making. 77% of working-age adults are in the agricultural sector, which makes up 28% of GDP. However, growth in the agricultural sector has grown slower than the national average (USAID, 2021). Women make up 52% of the labor share in agriculture (Chirsteansen, 2018), yet their adoption rate of technologies remains much lower than men's, particularly among married women. Because married women's adoption of technology or new crops generally requires spousal bargaining, understanding intra-household decision-making over adoption is critical to increasing agricultural production and women's empowerment (Theis et al., 2018). The households from the villages that we organize our experiments at are part of a climate-smart agriculture (CSA) project. In the project, households need to invest in new inputs and their labor to implement the CSA practices, therefore, the decision that couples need to make in our experiment are relevant for the study context.

In our experiment, couples were offered compensation to sort cups of beans in a forty-minute period. Couples were given an endowment and could use it to buy cups of unsorted beans, sort the cups (for a maximum of forty minutes), and return them for a monetary return, which depends on the type of beans they sorted. One type of beans offered a constant return of 67% if the cup was returned sorted, while another provided an uncertain return of either 0% or 300% upon return of the sorted cups. After forty minutes, cups left unsorted were not given any compensation, and a refund could not be given (i.e. -100% return). We randomized who in the couple performed the labor (men, women, or the couple could decide the person to sort) and the spouse making the purchase decision (man, women, or joint decision). The randomization of decision-makers and laborers allows for identifying the causal relationship between gender and decisions involving risk and labor costs (e.g. the adoption of a new crop). We compare the findings across groups and check for differences in investment behavior by the gender of decision-maker and laborer. After the experiment, couples were asked to (separately) complete a short exit survey to gauge their farm characteristics, demographics, risk preferences, bargaining power and labor costs within the household. We use these controls to provide robustness checks to our main statistical specifications.

Our findings suggest that who decides in and works for an agricultural investment within a household matter for the gender gap in agricultural investments. Men purchase the same amount of' safe cups' as women but purchase more' risky cups', showing that they have higher and riskier investment levels. However, men end up leaving more unsorted cups and do not complete the tasks they invest in or are assigned to by their spouses, decreasing the overall efficiency of the investments. When couples decide jointly, they purchase the same amount of total cups that men would invest in alone. However, their allocation between risky cups and safe cups (or their investment portfolio mix) is different. Couples deciding jointly purchase the same level of risky cups as women would alone and purchase more safe cups instead, showing that bargaining decreases the market risks the couples take in contrast to the riskier choices of man spouses. Further, compared to women, fewer cups overall are assigned to man laborers – a difference almost entirely explained through an increase in risky cups assigned to women. Still, men complete sorting their assigned cups less often. These differences in allocations could be because men appear to carry higher production risk, and assigning them risky cups could exacerbate the riskiness of the allocation.

These intra-household dynamics highlight the difficulties in closing the gender gap through technology and crop adoption in (predominately) man-headed households. Women invest less but are assigned more labor, which puts a burden on their already full schedule of agricultural and domestic duties. To understand why joint decision thereby household bargaining has a better investment performance than men invest alone, we use random assignment to laborer and explore whether women alone and household bargaining are more successful in detecting the potential low returns to labor and therefore make better investment decisions. Our results show that when a couple makes the investment decision alone, they do not take into account the labor performance of their spouse. However, when they jointly decide, they invest less when the spouse assigned to labor is not expected to perform well. This implies that involving women in decision making reminds men that they do not sort beans well in our experimental setting.

These results imply that, to improve the overall agricultural investment performance, programs must facilitate behavioral changes that help households in closing the gender gap. Those programs should empower women and involve women more in agricultural investment processes. To contribute to this policy discussion, we provide preliminary evidence on whether two women empowerment interventions, gender training and participation in VSLAs, which are both shown to have a positive effect on women empowerment (Bulte et al. 2016; Bulte et al. 2017, Karlan et al.,2017), are associated with the behavioral change in investment decisions. We specifically test whether investment decisions of couples get closer to women's preferences that are more efficient in our experimental context when women and men receive gender trainings or participate in (VSLAs).

Our households show that in households where women attended gender training in the past twelve months, women allocate more of their endowment to the high-return, risky investments as sole decision-makers and are more likely to jointly bargain with their husbands to reduce the production risks when making decisions together. This led to more profitable woman investment decisions and more efficient joint decisions. We also find that in the households where women participate in savings and loan associations (VSLAs) men's

investments are more in line with women's investment preferences, indicating that VSLAs have the potential to empower women (and men) in household investment decisions.
The rest of the paper is structured as follows. Section 2 gives an overview of the experiment design, setting, and process. Section 3 describes the data used in the evaluation, while Section 4 describes the empirical methodology used in comparing different outcomes between men, women, and joint decision-makers and laborers. We present the experiment results in Section 5 and give a discussion and conclusion in Section 6.

2. Experiment Design, Sample, and Process

2.1 Experiment Design

To analyze the role of gender in labor-intensive investment decisions, we conduct a lab-in-the-field experiment involving a real effort task. Real-effort tasks are non-trivial mental or physical tasks used in experimental settings to measure participants' labor response to a certain incentive structure (Lezzi et al., 2015). The real-effort task in our experiment involves sorting beans – a task employed in similar settings by Bulte et al. (2020). Green-colored beans were mixed with either grey-colored or red-colored beans[1] and placed in standard, disposable 65ml plastic cups, and couples were rewarded with cash for each cup of beans sorted in 40 minutes. To successfully sort a cup, they would need to place all green-colored beans in one cup and all of the grey/red-colored beans in another cup. Pre-testing showed that a fast bean-sorter could sort no more than 10 cups in 40 minutes.

Participants had to decide on how many cups of beans to sort before they began sorting. They were not allowed to change their decision after the bean-sorting began. Couples were given an endowment of 3,000 Tsh (about 1.29 USD). They could purchase a cup of unsorted beans for 300 Tsh (so they can by a maximum of 10 cups)[2]. Two types of cups were available to be purchased – cups of green-colored beans mixed with grey-colored beans (referred to as safe cups from here on) and cups of green-colored beans mixed with red-colored beans (referred to

1 All of the beans were the same size.

2 Respondents could not use their own cash to buy beans

as risky cups from here on). If a participant successfully sorted a safe cup, they were paid 500 Tsh (a return of 67% over the purchase price). If a participant successfully sorted a risky cup they would either be paid 900 Tsh (a 300% return) or 300 Tsh (a 0% return) with a 50% probability of each payment occurring. After sorting, the payment was determined by having participants draw a card from a full deck of playing cards. Drawing a red card indicated receiving the 0% return for all risky cups, and drawing a black card indicated receiving the 300% return for all risky cups. Participants could also choose not to purchase a cup to sort and keep their endowment. Therefore, participants could purchase any combination of safe and risky cups ranging from zero to ten cups in total. Unsorted cups were given a payment of 0 Tsh (a 100% loss) such that participants lost any endowment they put towards purchasing cups they do not sort.

To identify the role of gender in labor-intensive decision-making, we randomize the gender of the decision-maker (Treatment Arm 1) and the laborer (Treatment Arm 2) through a semi-factorial design. In the first treatment arm, couples were randomly assigned to have woman spouse, man spouse, or both members (joint) decide how many and which cups to purchase. In the second treatment arm, couples were randomly assigned to have men or woman perform the bean sorting. The semi-factorial design is shown in Table 1.

Table 1: Experiment Design

Treatment Arm1: Decision making	Treatment Arm 2		
	Man	Woman	Joint
Man	Man, Man	Man, Woman	NA
Woman	Woman, Man	Woman, Woman	NA
Joint	Joint, Man	Joint, Woman	Joint, Joint

The joint labor treatment indicates that couples can jointly decide whether men or women perform the bean sorting. This treatment provides a' real-world' scenario, as labor assignment is not exogenously determined in households but rather endogenously decided upon by household members. Since the experiment does not have a traditional control group, the joint decision-making and joint labor decision provide a logical comparison group for the other seven treatment groups.

In this experiment design, we are interested in seven main outcome variables: total number of cups purchased, the number of safe cups purchased, the number of risky cups

purchased, the number of cups left unsorted, the expected return of a decision, the standard deviation of a decision's returns, and the final payment issued to participants. The total number of cups purchased relates to production risk, while the number of risky and safe cups relates to market risk. The number of unsorted cups is an indicator of inefficiency (or realizations of production risk) in investment decision-making.

2.2 Sample

The experiment took place across fifteen villages in Iringa, Tanzania. In the villages, CARE-Tanzania implements Kukua ni Kujifunza (KnK) project that has introduced climate-smart agriculture villages and soybean value chain through farmers field schools. The project mainly targeted women farmers organized via farmer field schools. In addition to agronomic training, the project gave gender sensitization training. This training involves discussing wife's and husband's roles in the household and agricultural issues. The project also encouraged farmers to participate, save and invest in CSA practices (e.g., rhizobium inoculation, composting, soybean seeds, fertilizer) through village savings and loan associations in the villages. Therefore, increasing the investment of married women in agriculture, especially CSA practices, was a relevant topic in the study villages.

The experiment included 521 couples m involving the project while not excluding others from the villages . Village officials and officials recruited couples from CARE to come to village offices and participate in the experiment . There were two experiment sessions in each village: one in the morning and one in the afternoon. We aimed to have 18 couples participate in each session. If the quotas for a session were not met through recruitment (e.g. because of no-shows), then respondents were asked to contact additional couples to participate.

The 521 couples were required to be married and involved in small-scale agriculture. Both members of the couple were required to be between 18 and 65 years of age. Verification of the eligibility requirements was carried out by village officials and was verified by checks in the survey administered to each participant.

2.3 Experiment process

The experiment took place outside of village offices in each village. While the setup slightly varied based on the location of the village office, a standard approach was taken to ensure that decisions were made in isolation and participants who had yet to begin the experiment could not see the bean-sorting area (i.e. they could not see what other couples had decided). This

setup alleviated concerns over observability. Members of each couple were also separated at the beginning of each session so that they could not discuss the experiment before participating in it. A sample setup can be seen in Figure 1[3].

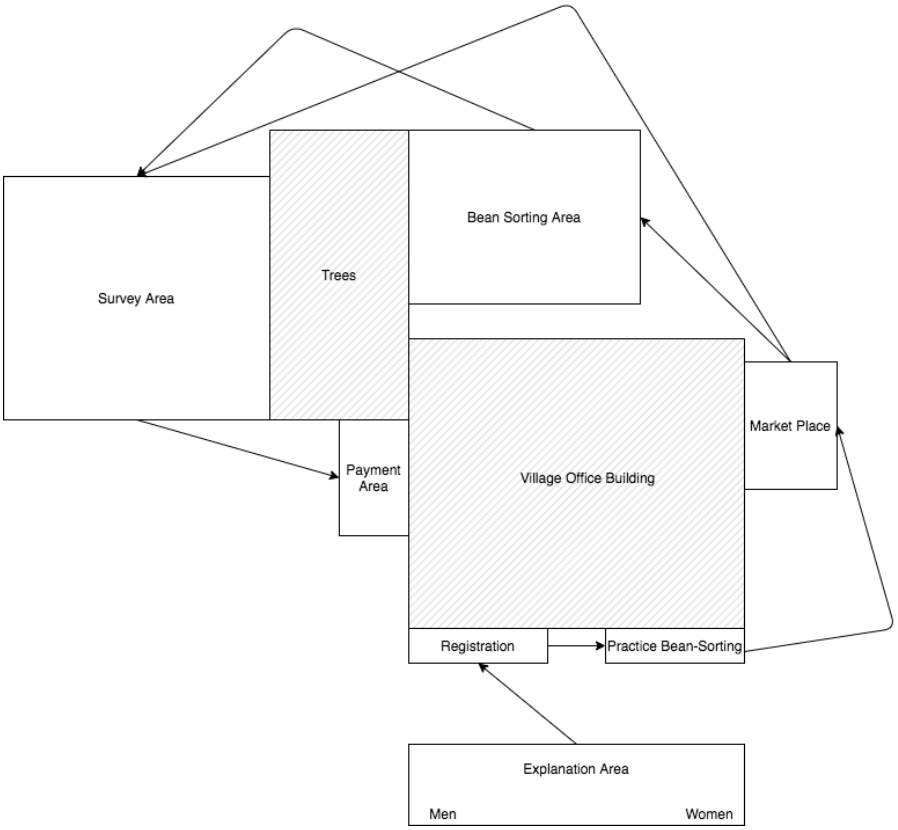

Figure 1: Example Experiment Setup

Each session lasted about four hours, but individual couples were compensated and able to leave once they completed the experiment. Once at least ten couples arrived at the beginning of a session, the experiment began with a detailed explanation of the experiment process to the entire group. Couples who arrived after the first explanation were given a subsequent explanation. Respondents were able to ask questions at any point during the explanation. During the explanation, men were asked to sit away from women so that couples could not discuss the experiment before making a decision.

After the explanation, couples were called one by one to register. At registration, they were required to give their names and ages and sign a consent form. Then both members of the

3 The arrows indicate the direction in which participants went through the experiment. From the marketplace, one member went to the bean-sorting area and another to the survey area based on which treatment group to which the couple was assigned.

couple practiced sorting one cup of beans in four minutes (1/10 of the time allotted for the bean-sorting in the experiment). This allowed the participants to understand their own and their spouse's ability in sorting beans. Before and after the practice bean-sorting, participants were asked questions about the experiment to test their understanding. If the participants did not have a firm understanding of the experiment, then the experimenters would explain the process again and then ask follow-up questions until the participants fully understood the process. A strict monitoring process was in place so that couples could not discuss the experiment among themselves during the registration or practice bean-sorting phases.

Participants then passed to the marketplace, where they were assigned to their treatment group by picking a slip with a treatment group from a bowl. Each couple only picked one slip. Before drawing a slip, the participants were once again asked about and could discuss their (mis)understanding of the experiment, particularly regarding the randomization process and decision options. After drawing a slip, the relevant participant(s) decided how many and which cups to purchase. If a sole decision-maker was assigned, the decision was made in the absence of their spouse. To make the decision, the experimenter handed the decision-maker(s) ten vouchers worth 300 Tsh each (for a total endowment of 3,000 Tsh). The decision-maker(s) could then spend (or keep) their vouchers as they wished. Non-observability was ensured such that other participants could not view the marketplace.

After the relevant participant(s) decided on how many and which cups to purchase, the cups were delivered to the bean-sorting area for the relevant participant to sort. The participant was given 40 minutes to sort the cups. Spouses who were not sorting beans were not allowed to help or observe their spouses sorting.

Finally, both members of each couple were required to complete a 10-minute exit survey before receiving payment. Once both members of each couple completed all of the tasks, then the couples could deliver their sorted and unsorted beans to the marketplace and receive payment based on what they sorted.

3. Data and variables

3.1. Investment outcomes

There are seven variables: the total number of cups purchased, the number of safe cups purchased, the number of risky cups purchased, the number of cups left unsorted, the expected return of the decision, the standard deviation of the decision's return, and the final

payment. These variables were recorded on paper forms when participants made their decisions (in the case of the variables related to cups purchased) and the bean-sorting and payment areas (for unsorted cups). All recorded figures were cross-checked when payments were distributed to participants at the end of the experiment. Based on the four measures relating directly to the number of cups, we calculate the expected return and standard deviation of decision portfolios and use these as additional outcomes along with the final payment distributed to each couple.

3.2. Independent Variables

The main independent variables are categorical variables for assignment to treatment groups. The levels of the categorical variables correspond to the groups defined in Table 1. When couples completed sorting beans in the experiment, we administered a short (less than ten minutes) exit survey to both men and woman in each couple. In the exit survey, we asked some questions to be used as control variables in our econometric specifications for robustness checks. Comparisons of these variables also give interesting insights into men and women's behavior.

Each participant was asked after the experiment which spouse is better at sorting. Respondents could report that men are better, women are better, or both spouses sort beans equally well. We combined these answers to create a couple-level indicator variable for which spouse is better at sorting. If both spouses report women (man) as the better sorter, we indicate that women (man) are better at sorting beans. If one spouse says women (man) is better and the other says they are equal, then women (man) are indicated by the couple-level variable to better sort beans. If both spouses said they were equal, then neither spouse was determined to better at sorting, and if both spouses reported themselves as the best sorter, neither spouse was reported to be better. Based on these calculations, we create an indicator variable, suboptimal, for whether the member of the couple who is relatively worse at sorting beans is assigned to be the laborer. This variable is equal to zero if the better sorter is assigned or if the spouses have an equal ability to sort. The variable is not defined for the couples assigned to decide which spouse sorts beans because this is an endogenous decision.

In terms of demographics, individuals reported their ages and education levels. We qualitatively observed that age was a potential correlate of the number of cups purchased during the

experiment. Household size may also be correlated with how spouses view labor costs, so we include this in the econometric specifications with controls.

Respondents were also asked about their farm characteristics and participation in local organizations. Respondents reported their farm size in hectares and listed the two most important crops on their farms. They were also asked if they earn any income from activities in farming. We then inquired about knowledge and participation in local organizations/programs related to agriculture: farmer field schools, NGO gender training, VSLAs, and NGO marketing training.

The next set of questions focused on spousal bargaining power related to the two most important crops (self-reported). We asked who is mostly responsible for decision-making with regards to each of the following: crop choice decisions, input decisions, and selling decisions. The respondents could reply with man, woman, joint, other man in the household, other women in the household, or others. We create an index of joint decision-making that simply adds the number of decisions that are made jointly between spouses. We then asked about income control concerning the two most important crops. We create a similar index equal to two if income from both crops is controlled jointly, one if only income from one crop is controlled jointly, and zero if no income is controlled jointly. Finally, we ask about labor divisions between men and woman in the household for the two main crops. We ask who is mainly responsible for land preparation, planting, weeding, harvesting, and selling. As with decision-making and income control, we create a simple index for the extent of joint labor tasks in the household.

To gauge individuals' risk preferences, we ask simple questions about individual's willingness to take risks (Vieider et al., 2015). We adopt the questions of (Vieider et al., 2015) to an agricultural setting. Individuals report on a scale of zero to ten of how willing they are to take risks in general and in certain situations (where 0 is completely unwilling and ten is completely willing). After asking about their general risk preferences, we asked about the willingness to take risks in farming in general, input decisions, new crop adoption, new technology adoption, and marketing.

Finally, we asked questions about the experiment itself. To gauge preferences towards spousal cooperation, we asked spouses who were not assigned to bean sorting if they would rather be helping their spouse or resting. We then asked whether spouses thought they or their spouses

were better at sorting and whether they believe their spouse gets more easily distracted than them.

3.3 Participant characteristics

Table 2 presents the participant characteristics by gender. Men reported larger farm areas than women by about 0.2 acres. The average household size has four members and 2.68 members work on the farm. However, the median number of farm laborers is 2, highlighting the importance of spousal bargaining in farm outcomes. 39% of households receive income from wage work. Nearly all households reported one of their two main crops as maize. The majority of respondents also responded that beans were one of the most important crops. However, women were more likely to list beans than men. A small percentage of households listed a variety of other crops as being one of the most important.

In terms of participation in local organizations, nearly all participants were aware of both a VLSA in their village and a Farmer Field School (FFS) . However, attendance differs among men and women. Women are more likely to attend gender training, marketing training, VSLA, and FFS.

In contrast to Vieider et al. (2015), we find no statistical difference between men and women with regards to risk aversion measures. Men and women do not report different levels of joint decision-making, joint labor, or income control. And with regards to the experiment, 71% of women say they sort better than their husband while only 61% of men say that their wife sorts better. But, when asked which member of the couple gets distracted more easily (i.e. is less able to concentrate in general), both genders tend to think that they get more distracted than their spouse.

Table 1: Participant Characteristics by Gender

Variable	Man	Woman	Diff
Farm Area	3.09	2.87	-0.219
	(0.07)	(0.08)	(0.109)
Household Size	5.08	5.12	0.047
	(0.09)	(0.08)	(0.124)
Household Farm Laborers	2.68	2.71	0.036
	(0.05)	(0.05)	(0.071)
Outside Income (Yes/No)	0.39	0.43	0.034
	(0.02)	(0.02)	(0.030)
Main Crop: Maize	0.99	0.98	-0.006

Variable	Man	Woman	Diff
	(0.00)	(0.01)	(0.007)
Main Crop: Beans	0.56	0.61	0.053
	(0.02)	(0.02)	(0.030)
Knowledge of VSLA	0.96	0.96	0.000
	(0.01)	(0.01)	(0.012)
Knowledge of FFS	0.90	0.92	0.013
	(0.01)	(0.01)	(0.018)
Attended NGO Trainings	0.64	0.74	0.096
	(0.02)	(0.02)	(0.028)
Attended VSLA	0.42	0.69	0.267
	(0.02)	(0.02)	(0.029)
Attended FFS	0.57	0.64	0.070
	(0.02)	(0.02)	(0.030)
Attend Marketing Training	0.50	0.54	0.047
	(0.02)	(0.02)	(0.031)
Risk Aversion: General	6.23	6.03	-0.203
	(0.12)	(0.13)	(0.176)
Risk Aversion: Farming in General	5.54	5.36	-0.177
	(0.12)	(0.13)	(0.178)
Risk Aversion: Inputs	4.69	4.69	-0.002
	(0.12)	(0.12)	(0.173)
Risk Aversion: Marketing	4.53	4.55	0.017
	(0.12)	(0.13)	(0.177)
Risk Aversion: Crop Adoption	4.65	4.65	0.002
	(0.13)	(0.13)	(0.182)
Risk Aversion: Tech. Adoption	5.02	4.96	-0.058
	(0.11)	(0.12)	(0.166)
Involved in Farm Management	0.98	0.98	0.006
	(0.01)	(0.01)	(0.009)
Joint Decision-Making Index	2.14	2.12	-0.021
	(0.06)	(0.06)	(0.079)
Joint Labor Index	0.92	0.92	-0.003
	(0.01)	(0.01)	(0.012)
Income Control Index	0.49	0.48	-0.009
	(0.02)	(0.02)	(0.025)
Woman Sorts Better	0.60	0.71	0.107
	(0.02)	(0.02)	(0.029)
Spouse More Distracted	0.28	0.26	-0.024
	(0.02)	(0.02)	(0.027)
N	531	531	1062
Desire to Help Spouse Sort	0.67	0.85	0.179***
	(0.03)	(0.02)	(0.037)
N	286	241	527

Significance levels: * < 10% ** < 5% *** < 1%. Standard errors in parentheses

4. Empirical methodology

4.1 Preliminary checks

Before testing differences in outcomes during the experiment, we conduct several tests to ensure that the randomization was carried out properly and ensured balanced samples across groups. We check for balance across the seven treatment groups and across those assigned to have the relatively less productive spouse sort. Those were not assigned to have the relatively less productive bean sorter perform labor.

To test for balance across the seven treatment groups, we perform iterative t-tests comparing each group across the covariates discussed above. The iterative t-tests first compare the mean covariate values of Treatment Group 1 to Treatment Groups 2 through 7, then Treatment Group 2 to Treatment Groups 3 through 7, and so on. Twenty-one comparisons are made for each variable using this method. We count the number of statistically significant differences found in all comparisons for each variable with the idea that around five percent of the differences should be statistically significant if using a 95% confidence test. The results of these comparisons are presented in Table A.A1 in the Appendix. Only farm area and the order in which the couple experimented within their session do not appear to be balanced across groups. Those assigned to joint decision-making for labor decisions appear to partake in the experiment later in each session. Since the research team did not know the area of any of the participant's farms and the assignment to each experimental group was completely random, these differences can only be explained as oddities arising from the randomization. Nevertheless, we control for all of the covariates in robustness checks to ensure that our analysis shows the true effect of the randomized assignment on experiment outcomes.

We then perform a similar check for balance across covariates for couples assigned to have the relatively less productive member sort beans and those assigned to have the relatively more productive member and equally productive spouse sort. We perform student t-tests across both groups along the covariates discussed earlier. The results are displayed in Table Appendix Table A.A2 and show that couples were successfully balanced across observables for these groups.

4.2 Estimation Strategy

4.2.1 Main results

The empirical analysis of the experiment relies on randomized assignment to identify the effects of gender and bargaining on outcome variables in the game. The main analysis first compares mean outcomes by the gender of the decision-maker and joint decision-making (Treatment Arm 1). The comparison is made specifically in three groups: Man spouse vs. woman spouse decision-making, man spouse vs. joint decision-making, and woman spouse vs. joint decision-making. To test these results, the following regression model is tested:

$$Y_i = \beta_0^d + \beta_1^d Woman\_decision_i + \beta_1^d Joint\_decision_i + \epsilon_i^d \quad (1)$$

where i denotes the couple, Y is a vector of outcomes, $Woman\_decision$ and $Joint\_decision_i$ equals 1 when women are the decision-maker (0 otherwise) and $Joint\_decision_i$ equals 1 when couple decides jointly (0 otherwise). ϵ is an idiosyncratic error term. The base of $Woman\_decision_i$ and $Joint\_decision_i$ is man decision-makers. β_1^d (β_2^d) estimates the difference between the investment outcomes of women (joint) investment decisions compared to the outcomes of men.

An analogous approach is used when estimating differences in means between couples randomly assigned different laborers (Treatment Arm 2). To test these differences, the following model is estimated:

$$Y_i = \beta_0^l + \beta_1^l Woman\_laborer_i + \beta_2^l Joint\_laborer_i + \epsilon_i \qquad (2)$$

where $Woman\_laborer$ equals 1 when women are assigned to sort (0 otherwise) and $Joint\_laborer_i$ equals 1 when the couple jointly decides who will sort (0 otherwise). ϵ is an idiosyncratic error term. The base category is the men is assigned to sort beans in the experiment. β_1^l (β_2^l) estimates the difference between the outcomes of women (joint) laborer compared to men investment decisions. Standard errors are estimated in the same manner as in Equation 1 above.

4.2.2 Analysis of the investment decision under sub-optimal labor allocation,

Why do investment decisions change when the household make investment decision jointly? To shed some light on this, we examine the change in investment decisions when a spouse with low returns to labor is assigned to the task. Our experiments randomize the spouse assigned for sorting beans, implying that spouses with low returns to labor are randomly assigned to sorting beans. We define those assignments as suboptimal allocation of household labor. Table A.A2 (in the Appendix) shows that those assigned to a 'sub-optimal' laborer and those assigned to an optimal laborer (i.e. the laborer with relatively high marginal returns) are balanced across the additional covariates.

Using the definition of suboptimal allocation, we first test whether suboptimal allocation affects the overall investment decisions of the couple. The model takes the form:

$$Y_i = \alpha_0 + \alpha_1 suboptimal_i + \epsilon_i \qquad (3)$$

where $suboptimal$ is an indicator variable for whether the couple is assigned to have the relatively less productive (i.e. sub-optimal) spouse sort beans. The base of $suboptimal$ is couples in which the relatively more productive spouse sorts or both spouse are equally productive. In this analysis, we exclude the couples assigned to the group in which couples decide which spouse sorts the beans are excluded from the estimation of Equation 3 (because those couples endogenously decide labor assignment). β_1 is the difference in outcomes between couples with sub-optimal labor allocations imposed on them and those who do not have such labor allocations imposed on them.

Next we test whether the women, men, or join investment decisions changes when there is a suboptimal labor allocation, using the following model:

$$Y_i = \alpha_0^d + \alpha_1^d suboptimal_i + \alpha_2^d Woman\_decision_i + \alpha_3^d Joint\_decision_i + \alpha_4^d Woman\_decision_i \times suboptimal_i + \alpha_5^d Joint\_decision_i \times suboptimal_i + \epsilon_i \quad (4)$$

From the model we test $\alpha_4^d \neq 0$ to investigate whether investment decision of women change and check $\alpha_5^d \neq 0$ to examine whether joint investment decision changes under suboptimal labor allocation.

Finally we also investigate the change in investment outcomes under suboptimal labor allocation.

$$Y_i = \alpha_0^l + \alpha_1^l suboptimal_i + \alpha_2^l Woman\_laborer_i + \alpha_3^l Woman\_laborer_i \times suboptimal_i + \epsilon_i \quad (5)$$

where we test $\alpha_3^l \neq 0$ to examine whether the investment decisions change when women are assigned as laborer and but returns to their labor is low – suboptimal.

4.2.3 Heterogeneity analysis

Can programs such as gender training or VSLAs change investment decisions? To test this tentatively, we investigate the heterogeneity in investment decisions by women's and men's participation in gender training or VSLAs in the past 12 months.

We test the relationship between participating gender training or VSLAs and investment outcomes using the following econometric specification. We estimate this model for participation in gender training and VSLAs separately

$$Y_i = \gamma_0^{d,g} + \gamma_1^{d,g} Women\_decision_i + \gamma_2^{d,g} Joint\_decision_i + \gamma_3^{d,g} participation_i^g + \gamma_4^{d,g} participation_i \times Women_{decision_i} + \gamma_5^{d,g} participation_i^g \times Joint\_decision_i + \epsilon_i \quad (6)$$

where g indicates whether women or men participated in the gender training (VSLA). $participation_i^g$ is an indicator variable equals 1 for whether men/women attended gender (VSLA) in the past twelve months (0 otherwise). We test $\gamma_3^{d,g} > 0$ to understand whether the participation is positively associated when men decide on the investment. We examine whether participation to training or VSLA changes the decision of the woman spouse and joint decision of the couple when compared to men, testing $\gamma_4^{d,g} \neq 0$ and $\gamma_5^{d,g} \neq 0$ respectively.

4.2.4 Estimation strategy and robustness check

To estimate the models above, we use Poisson models for our count data outcome variables: the total number of cups, number of risky cups, number of safe cups, or the number of unsorted cups because these are all count variables. For expected portfolio return, portfolio standard deviation, and final payment, OLS is used for the estimation of coefficients and standard errors. We do not cluster standard errors in the main analysis for analysis of differences in means of Treatment Arm 1 (decision-making), Treatment Arm 2 (labor assignment), and sub-optimal labor allocations. Since the treatment was assigned individually, and we do not attempt to make claims about the population as a whole, the standard errors should not be clustered (Robinson, 2020). For robustness, we cluster the standard errors at the treatment-session level. See Appendix B for these results.

Since the couples are allocated randomly to the treatment arms, we present the main results without control variables and fixed effects. To test for the robustness of the results from the models above, we condition the comparison of means on the control variables. These control variables also include, session and village fixed effects and the order in which the participants began the experiment in each session. We estimate the modes again using appropriate distributional assumptions. In the robustness checks, we cluster standard errors at the treatment group-session level to account for correlations between couples assigned to the same treatment group within the same session. The robust results with controls and fixed effects are presented in Appendix B.

5. Results

5.1 The 'Natural-Case': The Joint-Joint Group

The 'Joint-Joint' group refers to the randomization group that makes decisions jointly and decides which spouse sorts beans. This group is intended to provide a 'natural' case – the most similar scenario to the real world. Out of the 76 couples assigned to this group, only 20 (26%) chose men to sort, while 56 (74%) chose women to sort. These choices can reflect two factors. First, women tend to be better at sorting, so the optimal decision is to decide that women sort in many cases. 86% of the couples assigned to this group chose the partner who the spouses believe is better at sorting to sort. Second, this tendency to decide that women sort can reflect the fact that women tend to have higher labor burdens than men in agricultural and household duties (Palacios-Lopez, Christiaensen, & Kilic, 2017). Therefore, assigning labor costs to women is consistent with labor allocations in small-scale farming households.

Table 2 displays the experiment outcomes for the joint-joint group. We test for differences between couples who chose men to sort and those who chose women to sort, but due to the low sample size (particularly for man laborers), these results should be interpreted with extreme caution. The average couple chose to sort 5.86 cups with significant differences between the number of cups chosen for man and woman laborers. When men sorts, couples purchase 1.2 more cups. This difference is driven solely by an increase in investment in safe cups – the number of risky cups purchased is equivalent for man and woman laborers. Despite higher investment levels for man laborers, payment levels are not significantly different for man and woman laborers. While couples purchase more cups for man laborers, man laborers leave 1.2 cups unsorted on average – an amount equivalent to the difference in investment levels between man and woman laborers. This suggests that couples over-invest when choosing men to sort and could not realize the returns on this extra investment.

Table 2: Outcomes for Joint-Joint Decision-Makers

Variable	Chose Man	Chose Woman	Diff (F-M)	Total
Total Cups	6.75	5.54	-1.214*	5.86
	[0.74]	[0.41]	[0.812]	[0.36]
Safe Cups	4.6	3.41	-1.189**	3.72
	[0.93]	[0.41]	[0.886]	[0.39]
Risky Cups	2.15	2.12	-0.025	2.13
	[0.65]	[0.39]	[0.760]	[0.33]
Unsorted Cups	1.2	0.46	-0.736***	0.65
	[0.37]	[0.15]	[0.328]	[0.15]

Expected Return	4565	4319.64	-245.357	4384.21
	[166.11]	[105.82]	[203.078]	[89.70]
Investment Std. Dev.	0.21	0.21	-0.001	0.21
	[0.06]	[0.03]	[0.065]	[0.03]
Payment	4295	4057.14	-237.857	4119.74
	[304.74]	[136.32]	[291.283]	[127.98]
N	20	56	76	

Significance levels: * < 10% ** < 5% *** < 1%

Standard errors in brackets. For Total Cups, Grey Cups, Red Cups and Unsorted Cups, standard errors are calculated using a Poisson Distribution. For Investment Std. Deviation, Expected Return, and Payment, standard errors are calculated using a normal distribution, and significance levels are found using a student T-Test.

5.2 Outcomes by Decision-Maker

We now move away from the natural case and analyze what happens to experiment outcomes when labor and decision-making genders are imposed. Table 3 summarizes our key findings on how investment preferences change by the gender of the decision-maker in the household and under joint decision-making. Columns 2-4 of the Table report the average investment outcomes of the experiment - number of safe and risky cups invested, the corresponding expected return and variability (standard deviation) of their investment - when man and woman spouses decide alone or jointly. It also shows the average actual outcomes of those decisions regarding unsorted cups and actual return (payment) received by the spouses. Columns 5-7 report the differences in outcomes between man and woman spouses' decisions, man spouse and joint decision-making, and woman-spouse and joint decision-making using the estimates from model 2 without any control variables. The detailed econometric results in the Appendix show that our results are robust to controlling for the treatment arm's order, participant characteristics, and session effects. Finally, we also present detailed experiment outcomes at the treatment arm level in Figure 2 and refer to those to explain our key findings.

Men and woman spouses have different investment choices for total cups purchased, and the mix of cups purchase when they invest alone (Table 3). Man spouses solely choose about half a cup more than women on average (5.80 total cups vs 5.28 total cups). The difference in total cups picked is driven by increased investment in risky cups by men. Man and woman spouses deciding alone choose similar amounts of safe cups (2.96 and 3.03, respectively), but men choose about 0.63 more risky cups on average than women (2.84 vs 2.24). Figure 2 shows that average investment behavior does not change significantly by the gender of the spouse assigned to work as a result of the investment. Compared to women, men take more risks for the household, and as Table 3 shows, their decisions lead to higher expected returns and a high standard deviation investment portfolio.

Table 3: Outcome Variables by Decision-Maker

Variable	Woman	Man	Joint	Diff (M-W)	Diff (M-J)	Diff (W-J)
Total Cups	5.28	5.80	5.73	0.525*	0.07	-0.456*
	[0.25]	[0.27]	[0.21]	[0.27]	[0.25]	[0.25]
Safe Cups	3.03	2.96	3.48	-0.075	-0.523***	-0.448**
	[0.25]	[0.28]	[0.22]	[0.20]	[0.19]	[0.19]
Risky Cups	2.24	2.84	2.25	0.6***	0.592***	-0.007
	[0.24]	[0.27]	[0.19]	[0.19]	[0.17]	[0.16]
Unsorted Cups	0.59	0.82	0.69	0.239**	0.134	-0.105
	[0.11]	[0.13]	[0.09]	[0.10]	[0.09]	[0.08]
Expected Return	4280.27	4445.27	4372.12	164.998*	73.146	-91.852
	[64.46]	[71.96]	[51.76]	[96.64]	[86.49]	[82.59]
Investment Std. Dev.	0.23	0.27	0.23	0.042	0.04	-0.001
	[0.02]	[0.02]	[0.02]	[0.03]	[0.03]	[0.03]
Payment	3980.27	4194.59	4085.4	214.322	109.196	-105.126
	[95.10]	[118.85]	[78.53]	[152.34]	[136.61]	[123.95]
N	147	148	226	295	374	373

Column 2-4 presents average outcomes listed in the first row for man-spouse, woman-spouse, and joint decisions. Column 5-7 show the differences between those three groups. Significance levels: * < 10% ** < 5% *** < 1%. Standard errors are in parentheses. Standard errors for total cups, grey cups, red cups, and unsorted cups differences are estimated from Poisson distribution using the delta method. Standard errors for expected return and Investment Std. Dev. Differences are estimated from a standard normal distribution (student T-test).

When couples decide jointly through bargaining, spouses compromise by following men preferences for the number of cups (more cups) and woman preferences for a mix of cups (safe cups instead of risky cups). In joint couple decisions, the total number of cups (5.73) is statistically similar to man spouse decisions (5.80) and higher than the number of cups purchased by woman spouses (5.28). Jointly deciding, couples purchase at least 0.60 fewer risky cups and 0.52 more safe cups than men's purchase alone. Hence, men win the negotiation on the number of cups purchased, but women win the negotiation to buy the type of cups. This shows that couples compromise by following men's preferences for the investment quantity and the women's preferences for portfolio composition.

However, the high investment ambitions of man spouses do not lead to significantly better payoffs because they also result in more failed investments (i.e. cups that remain unsorted). The table shows that the actual payments from the experiment do not change significantly by the decision-maker. This is because about 0.82 cups, about 15% of the total cups purchased, are left unsorted when men spouse invests – 0.24 cups more unsorted cups than when women decide alone. Figure 2 shows that the high investment levels leading to unsorted cups are particularly observed in the man-man treatment arm (where the man spouse decides how many cups to purchase and is assigned to sort the cups). This implies that man spouses underestimate the cost of effort and their capacity to perform this labor-intensive task.

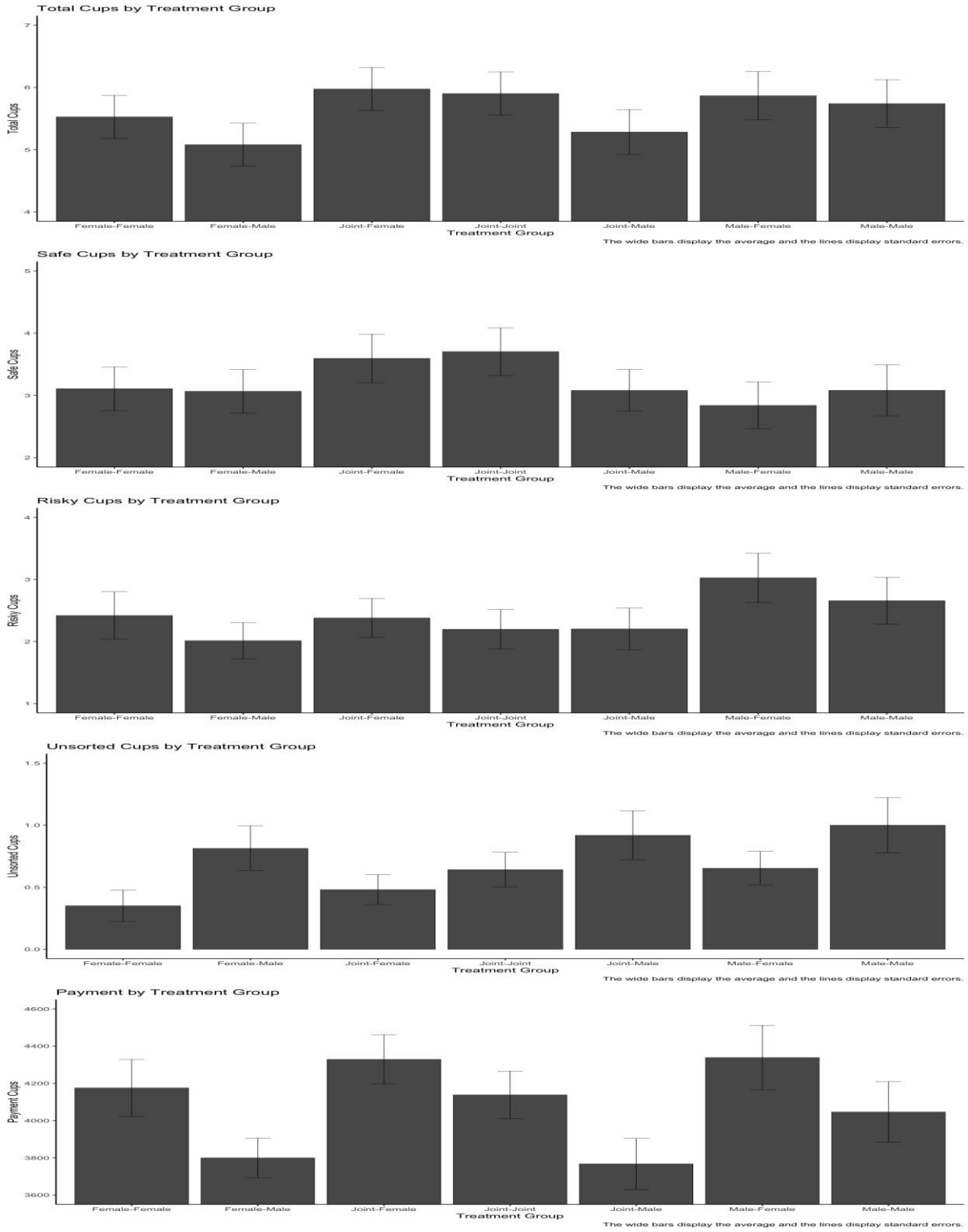

Figure 2: Experiment outcomes by treatment group

Notes to the Figure: In man-man treatment, man spouse purchase cups and he sorts. In man-woman treatment, man spouse purchases the cups and woman spouse sorts. In joint-man treatment, spouses jointly purchase cups and man spouse sorts. In joint-woman treatment, spouses jointly

5.3 Outcomes by Laborer

Now we turn to our results for investment preferences and outcomes by the gender of the spouse assigned to work. Columns 2-4 of Table 4 report the average outcomes, and Columns 5-7 of Table 4 show the investment differences between the labor assignment to man spouses, woman spouses, and a joint decision of which spouse works. This table estimates Model 2 without any control variables and compares the estimates between the different treatment arm assignments in Treatment Arm 2. The detailed econometric results (Tables A.B1-A.B7 in the Appendix) show that these results are robust to controlling for the treatment arm's order, participant characteristics, and village and session effects.

Table 4: Outcomes Variables by Laborer

Variable	Woman	Man	Joint	Diff (M-W)	Diff (M-J)	Diff (W-J)
Total Cups	5.79	5.37	5.86	-0.418*	-0.481	-0.063
	[0.21]	[0.21]	[0.36]	[0.22]	[0.32]	[0.32]
Safe Cups	3.18	3.05	3.72	-0.127	-0.669***	-0.542**
	[0.22]	[0.21]	[0.39]	[0.17]	[0.25]	[0.32]
Risky Cups	2.61	2.32	2.13	-0.291*	0.188	0.479**
	[0.21]	[0.20]	[0.33]	[0.15]	[0.20]	[0.20]
Unsorted Cups	0.5	0.92	0.66	0.418***	0.26**	-0.158
	[0.07]	[0.12]	[0.15]	[0.08]	[0.11]	[0.10]
Expected Return	4419.47	4306.85	4384.21	-112.62	-77.361	35.259
	[55.07]	[53.86]	[89.70]	[77.08]	[105.61]	[108.29]
Investment Std. Dev.	0.25	0.24	0.21	-0.018	0.023	0.041
	[0.02]	[0.02]	[0.03]	[0.024]	[0.03]	[0.03]
Payment	4286.73	3868.95	4119.74	-417.776***	-250.787	166.989
	[88.81]	[79.86]	[127.98]	[119.66]	[155.14]	[170.22]
N	226	219	76	445	295	302

Column 2-4 presents average outcomes listed in the first row for three groups: man-spouse , woman-spouse assigned to work, the spouse jointly decided to work. Column 5-7 show the differences between those three groups. Significance levels: * < 10% ** < 5% *** < 1%. Standard errors are in parentheses. Standard errors for total cups, grey cups, red cups, and unsorted cups differences are estimated from Poisson distribution using the delta method. Standard errors for expected return and Investment Std. Dev. differences are estimated from a standard normal distribution (student T-test).

We find differences in investment preferences and outcomes depending on the gender of the spouse assigned to perform the labor-intensive sorting task. Table 4 shows that man spouses are assigned to sort about 0.4 fewer total cups and 0.3 fewer risky cups than women spouses assigned to work. This decreases the expected return of the investment portfolio by about 112 Tsh. The detailed investment levels in Figure 2 show that this result is mainly driven by two treatment arms in which man spouses are assigned to sort beans: (i) women spouse decides alone on the cups to be sorted, and man spouse is assigned to sort; (ii) the couple decides jointly on the number of cups to be sorted, and man spouse is assigned to sort. In those treatments, women spouse's investment in risky cups is significantly less than the treatment arms in which the woman spouse is assigned to work, and she decides alone or jointly with man spouse on the number of cups to be invested. She does this by investing in fewer risky cups to decrease the overall risk. These results show that woman spouses expect that man spouses will not be good

at sorting beans. The exit survey result confirms this, reported that 60% of men and 71% of women said that women partner was better at sorting.

Women' perceptions of men are confirmed by the actual performance of men spouses in bean sorting. When work is assigned to man spouses, men leave about 0.9 unsorted cups, 0.4 more than women leave on average. Interestingly, the percentage of unsorted cups of total cups assigned to men is 16%-17% across all decision-makers and does not change by the number of total cups to be sorted. This shows that the bean sorting skills of man spouses do not explain the high number of unsorted cups left by men, and they do not do their best in the experiment. Instead, men perceive that their effort for this task is very costly, and they were perhaps bored of the task and left those cups unsorted. The exit survey findings are also in line with this interpretation of the result. The survey hypothetically asked spouses not assigned to sort beans whether they would desire to help their spouse in bean sorting to earn more money if they were allowed to help them. 67% of men answered 'yes' compared to 85% of women, showing that, although they would receive payment, fewer men prefer working in this labor-intensive job compared to women.

As a result of the lower initial investment in higher-paying risky cups and unsorted cups, when men spouse is assigned to be the laborer, the household's actual earning from the investment decreases. Households with man spouses assigned to sort beans earn about 418 Tsh less than the households with woman spouses assigned to work. Overall, these results show that the success of the investment in labor-intensive tasks depends on who will perform the job in the household and performs better when women are assigned to work due to man spouses' dislike of those labor-intensive tasks.

5.4 Bargaining Under Sub-Optimal Labor Allocations

Returns to labor differ within households, and households must make decisions for both laborers with relatively high returns and laborers with relatively low returns. By randomizing the laborer assigned to sort beans, we randomize whether the person with relatively low returns sorts. We define such an assignment as a 'sub-optimal labor allocation. Table A.A2 in Appendix A shows that those assigned to a 'sub-optimal laborer and those assigned to an optimal laborer (i.e., the laborer with relatively high marginal returns) are balanced across the additional covariates.

Table 5 shows the differences in means between sub-optimal and optimal laborers. Decision-makers assigned slightly fewer cups to sub-optimal laborers (0.8). Laborers with relatively low returns sort 0.44 fewer cups, and couples in these groups receive about 354 Tsh less than couples in which the optimal laborer is randomly assigned. These results are as expected – the

less productive member in couples performs worse and yields worse outcomes.

Table 5: Effect of Sub-Optimal Labor Allocation on Experiment Outcomes

	(1) Total Cups	(2) Risky Cups	(3) Safe Cups	(4) Unsorted Cups	(5) Investment Std. Dev.	(6) Payment
Sub-Optimal	-0.0856*	-0.106	-0.0694	0.444***	-0.0208	-353.6***
	(0.0444)	(0.0671)	(0.0592)	(0.115)	(0.0260)	(130.8)
Order Controls	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No
N	445	445	445	445	445	445

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

When given the task of choosing investment levels for a laborer with relatively low returns, do men, women, and joint decision-makers respond differently?
Table 6 shows that when men decide on a sub-optimal laborer, they do not change the total number of cups they invest in compared to when the optimal laborer is assigned (row 1). This leads to inefficient investment decisions – about 0.73 more cups are left unsorted, and payments are 455 Tsh lower when compared to cases of optimal labor allocation. Woman participants fair slightly better. They do not change their investment behavior either, but since their investment levels are already lower than men's, woman decision-makers avoid additional unsorted cups and losses in payment when the sub-optimal laborer is assigned.

Only the joint decision-makers change their behavior when a sub-optimal laborer is assigned. Joint decision-makers lower the total amount of cups that they invest in. This is driven by a reduction in investment in safe cups. As a result, joint decision-makers choose investment levels that result in 0.57 fewer unsorted cups. However, the reduction in investment levels and unsorted cups does not mean that joint decision-makers receive higher payments than individual decision-makers when a sub-optimal laborer is assigned. Table 7 shows that the results do not change by whether men or women are the sub-optimal laborer.

These results indicate that joint decision-making can lead to more optimal investment-labor allocations. When a laborer has relatively low returns, individual decision-makers assign labor as if they were assigning it to the more productive spouse. However, joint decision-makers reduce labor allocations to the less productive member. This indicates that spouses are discussing with each other and making more sensible decisions – they are 'talking each other down' from investment strategies that may not make sense.

Table 6: Effect of Sub-Optimal Labor Allocation on Experiment Outcomes by Decision-Makers

	(1) Total Cups	(2) Risky Cups	(3) Safe Cups	(4) Unsorted Cups	(5) Investment Std. Dev.	(6) Payment
Sub-Optimal	0.0256	-0.0947	0.136	0.729***	-0.0296	-454.8**
	(0.0743)	(0.109)	(0.102)	(0.181)	(0.0455)	(227.9)
Woman Decision-Maker	-0.0757	-0.235***	0.0657	-0.194	-0.0525	-339.5*
	(0.0591)	(0.0866)	(0.0812)	(0.185)	(0.0351)	(176.2)
Joint Decision-Maker	0.0553	-0.198**	0.263***	0.0694	-0.0290	-86.12
	(0.0571)	(0.0857)	(0.0776)	(0.173)	(0.0351)	(176.2)
Sub-Optimal x Woman Decision-Maker	-0.0645	-0.00529	-0.131	-0.352	0.0367	421.3
	(0.108)	(0.164)	(0.145)	(0.287)	(0.0643)	(322.5)
Sub-Optimal # Joint Decision-Maker	-0.270**	-0.0250	-0.471***	-0.567**	-0.00857	-105.9
	(0.108)	(0.161)	(0.145)	(0.273)	(0.0636)	(318.9)
Order Controls	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No
N	445	445	445	445	445	445

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

Table 7: Effect of Sub-Optimal Labor Allocation on Experiment Outcomes by Laborers

	(1) Total Cups	(2) Risky Cups	(3) Safe Cups	(4) Unsorted Cups	(5) Investment Std. Dev.	(6) Payment
Sub-Optimal	-0.0718	-0.0716	-0.0720	0.191	-0.0260	-227.7
	(0.0586)	(0.0892)	(0.0778)	(0.145)	(0.0345)	(172.2)
Woman Laborer	0.0336	0.0746	0.00136	-0.475***	0.000766	271.9*
	(0.0520)	(0.0786)	(0.0693)	(0.149)	(0.0312)	(155.8)
Sub-Optimal x Woman Laborer	0.0911	0.145	0.0449	-0.721	0.0767	586.9
	(0.146)	(0.213)	(0.199)	(0.603)	(0.0887)	(442.8)
Order Controls	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No
N	445	445	445	445	445	445

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

5.5 Heterogeneity analysis

5.5.1 Direct Empowerment Programs: Gender Training Participation

Can the investment behavior of women in the household bargaining process change via gender training? We give a tentative answer to this question, testing whether the investment decisions and outcomes from our experiment change in the households with women attending gender training organized by an NGO in the past twelve months.

Table 8 and Table 9 show the estimation results of Equation 5, testing the experimental investment outcomes of woman spouse and joint decision-making by the training attendance by men and women, respectively. The results indicate that women who attend gender training and are assigned the sole decision-making role in the experiment purchase similar amounts of cups to women who did not attend, but choose a different mix of cups. The estimate for the interaction term for woman decision-making and training shows that, on average, women who attend gender training choose 0.36 fewer safe cups and 0.31 more risky cups than those who do not– a nearly exact trade-off from the safe cups to the risky cups. These results show those woman spouses who participated in the training are more entrepreneurial and take more risks than those who do not. Further, in couples where women attended gender training, man decision-makers choose 0.23 fewer risky cups and 0.4 more safe cups, indicating that their decisions are more reflective of woman preferences than men in couples in which women did not attend gender training.

When men attend gender training, their investment levels are higher and riskier (similar to woman decision-making results when women attend). Men who attended gender training invested in 0.16 more cups, reduced allocations to safe cups, and increased allocations to risky cups by 0.71. This resulted in both higher expected return and higher payments when compared to man decision-makers who do not attend training. Women in couples in which men attended gender training do almost exactly the opposite – they reduce total investments by 0.22 cups, increase their allocation to safe cups by 0.29, and reduce investments in risky cups by 0.86 when compared to women in the couple in which men did not attend. Consequently, they have lower expected returns, lower portfolio standard deviations, and lower payments.

Table 8: Decision-Making Outcomes by Women's Gender Training Attendance

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Woman	-0.0974	0.280**	-0.456***	-0.0464	-224.1	-0.104*	-320.5
	(0.0981)	(0.141)	(0.141)	(0.281)	(180.7)	(0.0564)	(281.5)
Joint	0.184**	0.579***	-0.199*	0.258	163.2	-0.0550	112.6
	(0.0847)	(0.125)	(0.119)	(0.243)	(166.8)	(0.0521)	(259.9)
Woman attended Gender Training	0.0721	0.405***	-0.228**	0.318	14.64	-0.0617	123.5
	(0.0763)	(0.116)	(0.103)	(0.215)	(145.9)	(0.0456)	(227.3)
Woman # Woman attended Gender Training	-0.000222	-0.341**	0.311*	-0.400	78.19	0.0862	135.5
	(0.114)	(0.161)	(0.165)	(0.325)	(211.5)	(0.0661)	(329.5)
Joint # Woman attended Gender Training	-0.263***	-0.554***	-0.0259	-0.585**	-306.4	0.0241	-296.8
	(0.0991)	(0.142)	(0.143)	(0.280)	(194.2)	(0.0607)	(302.6)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$. Notes: Man spouse deciding on the investment levels is the base category.

However, this does not suggest that men attending gender training have a negative influence on the woman's investment decisions. On the contrary, women in couples in which men attended the training have more bargaining power than women in couples in which men did not attend. When making joint decisions, couples make decisions more in line with women's preferences when men attend gender training. Table 3 shows that overall, couples choose an investment level in line with man preferences and an investment mix in line with woman preferences when jointly bargaining. However, if men attend gender training, women gain more. Table 9 shows that when men attend gender training, couples jointly deciding invest in fewer total cups and fewer risky cups. This shows that women are more successfully negotiating their preferences if men attend gender training than those who do not. These negotiation results are even more favorable to woman spouses than when only women attend gender training, indicating that including men in gender training is imperative to improving women's bargaining power.

Table 9: Decision-Making Outcomes by Men's Gender Training Attendance

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Woman Decision-Maker	0.0508	-0.151	0.367***	-0.530**	130.5	0.0684	261.0
	(0.0862)	(0.111)	(0.139)	(0.253)	(160.3)	(0.0496)	(249.1)
Joint Decision-	0.118	0.125	0.105	-0.230	149.6	0.0200	31.46

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Maker	(0.0749)	(0.0914)	(0.131)	(0.199)	(140.5)	(0.0435)	(218.3)
Man attended Gender Training	0.164**	-0.271***	0.705***	-0.0464	361.5***	0.144***	388.4*
	(0.0731)	(0.0967)	(0.118)	(0.188)	(137.9)	(0.0427)	(214.3)
Woman # Man attended Gender Training	-0.220**	0.287**	-0.856***	0.272	-455.8**	-0.170***	-723.9**
	(0.105)	(0.140)	(0.165)	(0.305)	(197.6)	(0.0612)	(307.0)
Joint # Man attended Gender Training	-0.200**	0.0569	-0.448***	0.0837	-346.5*	-0.0921*	-213.0
	(0.0927)	(0.121)	(0.152)	(0.250)	(176.6)	(0.0547)	(274.3)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10 reports the estimation results for the investment outcomes by laborer and woman spouse's participation in gender training. We focus on the interaction terms (Woman × training and Joint × training), testing to test the difference in numbers of cups assigned to women compared to the number of cups assigned to men when women participated in the training. We also check whether those differences influence the investment outcomes. Our findings show that women's participation in gender training does not play a major role in the number of cups assigned to sort. We only find that the interaction term for women assigned to work and training is statistically significant for unsorted cups, despite not being assigned fewer cups. This could reflect a difference in the motivation of attendees and non-attendees rather than the influence of training on labor outcomes.

Table 10: Labor Outcomes by Women's Gender Training Attendance

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Woman Laborer	0.103	0.0980	0.108	-0.0597	143.3	0.00678	474.3*
	(0.0833)	(0.115)	(0.121)	(0.240)	(159.5)	(0.0498)	(246.5)
Joint Laborer	0.277***	0.298**	0.253*	0.163	411.2**	0.0431	372.5
	(0.0941)	(0.129)	(0.138)	(0.268)	(192.5)	(0.0601)	(297.4)
Woman attended Gender Training	0.0177	0.129	-0.122	0.352**	-9.950	-0.0237	41.17
	(0.0640)	(0.0870)	(0.0949)	(0.168)	(119.2)	(0.0372)	(184.2)
Woman Laborer x Woman attended Gender Training	-0.0365	-0.0862	0.0308	-0.719***	-36.32	0.0168	-75.02
	(0.0953)	(0.130)	(0.140)	(0.275)	(182.2)	(0.0569)	(281.5)

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Joint Laborer x Woman attended Gender Training	-0.291**	-0.140	-0.577***	-0.720**	-488.1**	-0.0969	-176.7
	(0.117)	(0.154)	(0.184)	(0.334)	(231.9)	(0.0724)	(358.2)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

Table 11 shows the results of the experiment for when men attend gender training. The coefficient for Woman x Man Attended Gender Training on unsorted cups is positive (but insignificant), indicating that women and men in couples in which men attended gender training leave similar amounts of cups unsorted. As in Table 10 above, this could reflect higher motivation levels for man participants.

Table 11: Labor Outcomes by Men's Gender Training Attendance

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Woman	0.0322	0.0590	-0.0142	-0.849***	33.28	-0.000968	527.5***
	(0.0672)	(0.0844)	(0.111)	(0.201)	(128.8)	(0.0400)	(198.1)
Joint	-0.0334	0.0452	-0.181	-0.828***	-70.46	-0.0260	242.1
	(0.0940)	(0.115)	(0.162)	(0.300)	(176.5)	(0.0548)	(271.5)
Man attended Gender Training	-0.0430	-0.198**	0.172*	-0.177	-7.646	0.0427	170.1
	(0.0608)	(0.0792)	(0.0960)	(0.145)	(114.8)	(0.0356)	(176.6)
Woman # Man attended Gender Training	0.0660	-0.0368	0.194	0.374	124.3	0.0306	-167.9
	(0.0838)	(0.109)	(0.133)	(0.248)	(160.3)	(0.0498)	(246.6)
Joint # Man attended Gender Training	0.187	0.243*	0.151	0.740**	238.6	0.00761	23.54
	(0.117)	(0.146)	(0.195)	(0.353)	(222.9)	(0.0692)	(342.7)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

5.5.2 Indirect Empowerment Programs: VSLA Participation

Does participation in programs that can indirectly enhance agency have similar associations with experiment outcomes as gender training? A local NGO facilitated VSLA programs, and

respondents were asked if they attended a VSLA in the past years. Table 12 shows the results of the experiment by VSLA attendance of woman participants and randomized decision-making groups. When women attend the VSLA, their decision-making in the program is not significantly different (please see the coefficients for Woman Decision-Maker x Woman attended VSLA). Similarly, joint decision-making does not differ between couples in which women attended and couples in which women did not attend. However, man decisions are different between the two groups. Men in couples in which women attended a VSLA invest in similar amounts of total cups but have different investment allocations. Men in couples of woman attendees invest in 0.24 fewer risky cups and 0.27 more safe cups – a near-perfect trade-off. These portfolio mixes are more in line with woman preferences. This suggests that woman attendance to VSLAs facilitates women influencing their partner's preferences – a sign of empowerment.

Table 12: Decision-Making Outcomes by Women's VSLA Attendance

	(1) Total Cups	(2) Safe Cups	(3) Risky Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Woman	-0.137	0.0892	-0.338***	-0.486*	-244.7	-0.0524	-2.683
	(0.0905)	(0.133)	(0.125)	(0.270)	(170.6)	(0.0531)	(264.3)
Joint	-0.0302	0.264**	-0.312***	-0.180	-124.6	-0.0411	-70.94
	(0.0825)	(0.121)	(0.115)	(0.231)	(159.7)	(0.0497)	(247.3)
Woman attended VSLA	0.00562	0.272**	-0.241**	0.0748	-65.58	-0.0313	141.0
	(0.0763)	(0.114)	(0.104)	(0.206)	(148.7)	(0.0463)	(230.4)
Woman # Woman attended VSLA	0.0638	-0.0677	0.138	0.212	114.5	0.0134	-307.3
	(0.108)	(0.154)	(0.155)	(0.316)	(204.9)	(0.0638)	(317.4)
Joint # Woman attended VSLA	0.0215	-0.134	0.107	0.0123	65.41	-0.00121	-62.70
	(0.0976)	(0.139)	(0.140)	(0.271)	(189.3)	(0.0590)	(293.2)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	520	520	520	520	520	520	520

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

Table 13 displays the results for random assignment to man, woman, or joint decision-making by men's VSLA attendance. Compared to man decision-makers who did not attend VSLAs, those that did attend the VSLAs make investment decisions that result in fewer unsorted cups, despite having similar investment levels. This suggests that the intrinsic motivation of attending a VSLA plays a role in the number of cups left unsorted. For women, decision-making is not significantly different in any outcomes between couples in which men attended and couples in which men did not attend. When making joint decisions, couples in which men attended a VSLA invest in 0.27 more risky cups than couples in which men do not attend.

This offsets the negotiating pattern seen earlier – men who attend a VSLA win the negotiation on both investment and portfolio mix levels. Perhaps this is due to increased bargaining power from men because they are involved in more investment-related activities.

Table 13: Decision-Making Outcomes by Men's VSLA Attendance

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Woman	-0.135**	-0.0422	-0.243**	-0.498***	-213.3	-0.0247	-277.3
	(0.0678)	(0.0921)	(0.100)	(0.188)	(129.7)	(0.0404)	(201.3)
Joint	-0.0550	0.152**	-0.337***	-0.356**	-147.1	-0.0478	-259.7
	(0.0578)	(0.0776)	(0.0879)	(0.153)	(113.4)	(0.0353)	(176.0)
Man attended VSLA	-0.0679	-0.0940	-0.0409	-0.367**	-90.24	0.0102	-166.4
	(0.0686)	(0.0962)	(0.0979)	(0.187)	(133.4)	(0.0415)	(207.1)
Woman # Man attended VSLA	0.0859	0.144	0.0148	0.369	103.4	-0.0355	135.8
	(0.0993)	(0.135)	(0.147)	(0.285)	(188.9)	(0.0588)	(293.2)
Joint # Man attended VSLA	0.102	-0.00469	0.271**	0.445*	183.3	0.0266	369.9
	(0.0896)	(0.122)	(0.133)	(0.249)	(174.5)	(0.0543)	(270.8)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	520	520	520	520	520	520	520

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1l * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

In terms of different decision-making by the assigned laborer, woman VSLA attendance does not play a major role. The coefficient for Woman Laborer x Woman Attended VSLA suggests that women who attended a VSLA have similar outcomes to those who did not attend when women are assigned the laborer role. Similarly, the coefficient for Joint Laborer x Woman Attended VSLA shows that the number of total cups and grey cups is lower compared to man laborers in couples in which women did not attend a VSLA, but compared to couples who did not attend a VSLA and were assigned to the joint laborer role, there is no difference in investment decisions (see the coefficients for Joint Laborer). Only couples with woman VSLA attendees leave fewer unsorted cups when assigned to the joint laborer group than those who did not attend a VSLA and were assigned to this group. As with unsorted cups for gender training attendees, these results could simply indicate that VSLA attendees have more intrinsic motivation than non-VSLA attendees.

Table 14: Labor Outcomes by Women's VSLA Attendance

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment

Woman	0.0935	0.140	0.0504	-0.439*	113.7	-0.00293	663.0***
	(0.0738)	(0.106)	(0.103)	(0.232)	(138.3)	(0.0431)	(211.7)
Joint	0.230**	0.564***	-0.209	0.258	212.2	-0.0294	167.9
	(0.0987)	(0.128)	(0.161)	(0.263)	(196.7)	(0.0613)	(301.2)
Woman attended VSLA	0.0797	0.342***	-0.233**	0.380**	28.71	-0.0413	145.0
	(0.0625)	(0.0875)	(0.0909)	(0.161)	(115.7)	(0.0361)	(177.2)
Woman # Woman attended VSLA	-0.0314	-0.152	0.119	-0.245	-3.733	0.0321	-353.1
	(0.0880)	(0.123)	(0.127)	(0.269)	(166.4)	(0.0519)	(254.8)
Joint # Woman attended VSLA	-0.224*	-0.544***	0.192	-0.854***	-209.8	0.00868	68.85
	(0.120)	(0.155)	(0.195)	(0.331)	(235.8)	(0.0735)	(361.1)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	520	520	520	520	520	520	520

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p <$ 0.10, ** $p < 0.05$, *** $p < 0.01$

When men attend VSLA's, the labor outcomes shift slightly. Compared to men who did not attend, those who attended VSLAs leave nearly 0.26 fewer cups left unsorted, once again reflecting the motivation of attendees and non-attendees. For woman laborers, there is no statistically significant difference in any outcome between couples in which men attended VSLA training and those who did not. However, for joint labor decision-makers (the natural case), VSLA attendance plays a role. Couples in which men attended invest in 0.38 more total cups than those who did not attend when assigned the joint-laborer role. This difference is driven by an increase in risky cups' investment and results in 0.75 more unsorted cups. Nevertheless, payments are 756 Tsh higher for couples assigned to the natural case when men attend VSLAs compared to those assigned the natural case when men do not attend. While man attendance to VSLAs may increase investment levels and returns, it does not increase women's bargaining power. Because woman spouses tend to make decisions that help improve household welfare overall more often than men, these higher investment levels in the real world may not translate to higher household welfare.

Table 15: Labor Outcomes by Men's VSLA Attendance

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Woman	0.0748	0.104	0.0376	-0.693***	95.24	-0.00464	504.8***
	(0.0527)	(0.0705)	(0.0792)	(0.153)	(101.1)	(0.0315)	(155.9)
Joint	-0.0654	0.235***	-0.622***	-0.629***	-184.9	-0.102**	-19.54
	(0.0740)	(0.0897)	(0.137)	(0.215)	(136.1)	(0.0425)	(209.9)
Man attended VSLA	-0.0623	0.0177	-0.170*	-0.258*	-105.4	-0.0418	3.507
	(0.0593)	(0.0781)	(0.0912)	(0.147)	(109.7)	(0.0342)	(169.1)

Woman # Man attended VSLA	0.00259	-0.149	0.197	0.222	43.53	0.0525	-196.7
	(0.0815)	(0.109)	(0.123)	(0.240)	(153.6)	(0.0479)	(236.8)
Joint # Man attended VSLA	0.382***	-0.114	1.171***	0.753**	723.9***	0.219***	756.0**
	(0.113)	(0.149)	(0.185)	(0.319)	(222.1)	(0.0693)	(342.3)
Order Controls	No	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No	No
N	520	520	520	520	520	520	520

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6. Conclusions

The gender gap in agriculture reflects pervasive bias, discrimination and denial of rights seen in other sectors of societies and economies. It also impedes development goals, and differences in adopting new technologies and crops by gender perpetuate the gap. Yet, little research has been done on spouses' behaviors when making investment decisions solely or jointly when labor costs are present. Understanding their behavior can help governments and international organizations bridge the gender gap in technology and new crop adoption more effectively. We analyze the behavior of small-scale farming couples making small investment decisions in a lab-in-the-field experiment with a real-effort task in rural Tanzania.

We find that man spouses are more likely to make larger investments with riskier returns. On the other hand, women are more efficient in their investments by investing at lower levels and minimizing 'failed investments'. When households jointly decide on the investment, they choose investments at quantities men would decide alone and select investment compositions in line with women's preferences. Our detailed analysis shows that this happens because when couples decide jointly, they consider that the couple assigned to the task might not perform well in the task (mostly men in our experimental context) they adjust their investments accordingly. Therefore joint decision making produces better investment outcomes. We also show that investment preferences and investment outcomes vary depending on which spouse is assigned to work on this labor-intensive task. When husbands are supposed to work, wives invest less because their husbands make less effort and do not complete the full task. This causes low performance in labor-intensive tasks (e.g., real effort experiment) among men and leads to investment failures and loss of money.

Our heterogeneity analysis shows that woman spouses who attended gender training in the past twelve months chose riskier investments and invested more. Moreover, when women who attended gender training bargain with their husband on the investment, couples decrease

the total investment levels in line with women's preferences. We found that, , men's investment decision is similar to women's in the households that women participated in VSLAs. On the contrary, in the households where men attended VSLAs men followed their risky preferences even stronger than men did not attend. This shows that there is a relationship between the bargaining power of spouses and participating in VSLAs.

We have two main policy conclusions on scaling small-scale farmer adoption of new labour-intensive agricultural practices requiring small financial investments (e.g., intercropping, manuring and composting). First, our results show that joint decisions on investment can lead to higher and safer investments, and promoting such decision-making could lead to higher and more stable incomes. Second, when invested in the tasks (or plots) that woman farmers are responsible for in the household, those joint investments lead to efficient outcomes thanks to hardworking women. This implies that development practitioners can support women's self-empowerment through improved bargaining power and promote joint decision-making on investment decisions within the family. Improved bargaining power will close the gender gap in the implementation of agricultural practices among married women in rural areas. Our tentative results on gender training show that especially gender training and, to some extent, VSLAs might be instruments to women empowerment and the bargaining power in joint decision-making.

Our paper also raises questions for new future research. First, it would be interesting to rigorously test the influence of gender training and VSLA's on women's bargaining power in agricultural investment and explore the mechanism. Future research should also investigate the barriers, drivers and incentives for joint decision making in the household investment decision that is more efficient than a decision than a man spouse make alone. Finally, our results also show that men in the household do not perform well with this labor-intensive bean sorting task. Therefore future research needed on how to change men's overvaluation of their own labour and undervaluation of women's labor, and how to improve joint decision-making within the household.

References

Adegbite, O. O. and Machethe, C. L. (2020). Bridging the financial inclusion gender gap in small-scale agriculture in Nigeria: An untapped potential for sustainable development. *World Development*, 127:104755.

Ali, D., Bowen, D., Deininger, K., and Duponchel, M. (2016). Investigating the Gender Gap in Agricultural Productivity: Evidence from Uganda. *World Development*, 87(May):152–170.

Buvinic, M., and Gupta, G. R. (1997). Woman-headed households and woman-maintained families: Are they worth targeting to reduce poverty in developing countries? Economic Development and Cultural Change, 45(2), 259–280.

Bulte, E., Lensink, R. and Vu, N., 2016. Gender training and female empowerment: Experimental evidence from Vietnam. Economics Letters, 145, pp.117-119.

Bulte, E., Lensink, R. and Vu, N., 2017. Do gender and business trainings affect business outcomes? Experimental evidence from Vietnam. Management Science, 63(9), pp.2885-2902.

Bulte, E., List, J. A., and van Soest, D. (2020). Toward an Understanding of the Welfare Effects of Nudges: Evidence from a Field Experiment in the Workplace. *The Economic Journal*, 130(632):2329–2353.

CARE (2021) A Win-Win for Gender and Nutrition Testing A Gender-Transformative Approach From Asia In Africa, http://careevaluations.org/wp-content/uploads/Win-Win-Final-Policy_Brief.pdf

Doss, C. R. (2001). Designing agricultural technology for African women farmers: Lessons from 25 years of experience. *World Development*, 29(12):2075–2092.

Doss, C., Meinzen-Dick, R., Quisumbing, A. and Theis, S., 2018. Women in agriculture: Four myths. Global food security, 16, pp.69-74.

FAO (2013). CEDAW – Convention on the Elimination of All Forms of Discrimination against Women. A tool for gender-sensitive agriculture and rural development policy and programme formulation.

Fisher, M. and Kandiwa, V. (2014). Can agricultural input subsidies reduce the gender gap in modern maize adoption? Evidence from Malawi. *Food Policy*, 45:101–111.

Gilbert, R. A., Sakala, W. D., and Benson, T. D. (2002). Gender analysis of a nationwide cropping system trial survey in Malawi. *African Studies Quarterly*, 6(1-2).

Goh, A. H. X. (2012). A Literature Review of the Gender-differentiated Impacts of Climate Change On Women's and Men's Assets and Well-being in Developing Countries. *International Food Policy Research Institute*, CAPRi Work(106).

Heise, L., Greene, M. E., Opper, N., Stavropoulou, M., Harper, C., Nascimento, M., Zewdie, D., Darmstadt, G. L., Greene, M. E., Hawkes, S., Henry, S., Heymann, J., Klugman, J., Levine, R., Raj, A., and Rao Gupta, G. (2019). Gender inequality and restrictive gender norms: framing the challenges to health. *The Lancet*, 393(10189):2440– 2454.

Huyer, S. (2016). Closing the Gender Gap in Agriculture. *Gender, Technology and Development*, 20(2):105–116.

International Labour Organization (2021). ILOSTAT database.

Karamba, R. W. and Winters, P. C. (2015). Gender and agricultural productivity: Implications of the Farm Input Subsidy Program in Malawi. *Agricultural Economics (United Kingdom)*, 46(3):357–374.

Kawarazuka, N., G. Prain, L. Forsythe, S. Mayanja, N.N. Mudege, C. Babini, and V. Polar. 2018. Gender in Agricultural Mechanization: Key guiding questions. Lima, Peru: International Potato Center.

Lezzi, E., Fleming, P., and Zizzo, D. J. (2015). Does it Matter Which Effort Task You Use? A Comparison of Four Effort Tasks When Agents Compete for a Prize. *SSRN Electronic Journal*.

Manfre, C., Rubin, D., Allen, A., and Summerfield, G. (2013). REDUCING THE GENDER GAP IN AGRICULTURAL EXTENSION AND ADVISORY SERVICES How to Find the Best Fit for Men and Women Farmers. (May).

Meinzen-dick, A. R. Q. R., Raney, T. L., and Croppenstedt, A. (2014). *Gender in Agriculture*.

Ngoma-Kasanda, E. and Sichilima, T., 2016. Gender and Decision Making in agriculture: A Case Study of the Smallholder Groundnuts Sector in Zambia. Lusaka, ZM: Musika Development Initiatives.

Njuki, J., Eissler, S., Malapit, H., Meinzen-Dick, R., Bryan, E. and Quisumbing, A., A review of evidence on gender equality, women's empowerment, and food systems.

https://bonndoc.ulb.uni-bonn.de/xmlui/bitstream/handle/20.500.11811/9132/fss_briefs_review_evidence_gender_equality.pdf?sequence=3&isAllowed=y

Nyasimi, M. and Huyer, S. (2017). Background: the gender gap in agriculture under a changing climate. *Agriculture for Development*, 30:2015–2018.

Palacios-Lopez, A., Christiaensen, L., and Kilic, T. (2017). *Women's Work on African Farms*.

Palacios-Lopez, A. and Lopez, R. (2014). Gender Differences in Agricultural Productivity : The Role of Market Imperfections. *Wp 14-01*.

Pauw, K. and Thurlow, J. (2011). Agricultural growth, poverty, and nutrition in Tanzania. *Food Policy*, 36(6):795–804.

Rola-Rubzen, M.F., Paris, T., Hawkins, J. and Sapkota, B., 2020. Improving gender participation in agricultural technology adoption in Asia: from rhetoric to practical action. Applied Economic Perspectives and Policy, 42(1), pp.113-125.

Robinson, T. (2020). When should we cluster experimental standard errors?.

Shibata, R., Cardey, S. and Dorward, P., 2020. Gendered intra-household decision-making dynamics in agricultural innovation processes: assets, norms and bargaining power. Journal of International Development, 32(7), pp.1101-1125.

Theis, S., Lefore, N., Meinzen-Dick, R., and Bryan, E. (2018). What happens after technology adoption? Gendered aspects of small-scale irrigation technologies in Ethiopia, Ghana, and Tanzania. *Agriculture and Human Values*, 35(3):671–684.

Udry, C. (2010). The economics of agriculture in Africa: Notes toward a research program. *AfJARE*, 5(1):284–299.

UNDP (2015). The cost of the gender gap in agricultural productivity. page 29.

USAID (2021). Economic Growth and Trade: Tanzania.

Van Aelst, K. and Holvoet, N. (2016). Intersections of Gender and Marital Status in Accessing Climate Change Adaptation: Evidence from Rural Tanzania. *World Development*, 79(July 2015):40–50.

van den Bold, M., Quisumbing, A. R., and Gillespie, S. (2013). Women's Empowerment and Nutrition: An Evidence Review. *SSRN Electronic Journal*, (October).

Vieider, F. M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., and Martinsson, P. (2015). Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries. *Journal of the European Economic Association*, 13(3):421–452.

Yokying, P. and Lambrecht, I. (2020). Landownership and the gender gap in agriculture: Insights from northern Ghana. *Land Use Policy*, 99(January):105012.

Appendix A: Checks for balance

Table A.A1: Balance along covariates across treatment groups

	M-M	M-F	F-M	F-F	J-M	F-F	J-J	Percent of Comparisons with Significant Differences
Male age	41.62	44.49	43.63	45.34	44.52	44.19	46.13	0.05
Female age	36.78	39.23	37.96	40.89	38.60	39.55	40.84	0.10
Household size	4.92	5.07	5.34	5.41	5.23	4.97	4.99	0.00
Man's number of spouses	1.04	1.11	1.07	1.11	1.04	1.03	1.04	0.10
Farm area	2.51	2.86	2.96	3.43	2.95	3.32	2.96	0.24
Wage income	0.68	0.64	0.52	0.54	0.63	0.62	0.64	0.05
Any spouse attended FFS	0.79	0.77	0.88	0.84	0.84	0.90	0.76	0.10
Any spouse attended gender training	0.88	0.89	0.89	0.92	0.89	0.92	0.87	0.00
Any spouse attended VSLA	0.77	0.85	0.79	0.78	0.71	0.79	0.80	0.05
Join decision-making index	2.18	1.88	2.04	2.09	2.32	2.19	2.22	0.10
Labor index	0.93	0.92	0.93	0.93	0.93	0.91	0.92	0.00
Income control index	0.44	0.47	0.51	0.52	0.50	0.46	0.47	0.00
Distraction treatment	0.44	0.48	0.45	0.50	0.47	0.49	0.42	0.00
Experiment order	9.22	10.87	9.41	10.18	9.88	10.44	13.34	0.29
Observations	73	75	73	74	73	77	76	

Table A.A2: Balance across covariates by assignment of optimal or sub-optimal laborer

	Optimal Laborer	Sub-Optimal Laborer	Difference
Male age	43.82	44.33	0.517
	[0.68]	[1.01]	[1.231]
Female Age	38.86	38.81	-0.057
	[0.64]	[0.91]	[1.142]
Household Size	5.15	5.18	0.033
	[0.11]	[0.16]	[0.197]
Man's number of spouses	1.06	1.08	0.023
	[0.01]	[0.02]	[0.025]
Farm area	3.03	2.96	-0.062
	[0.10]	[0.11]	[0.168]
Wage income (y/n)	0.59	0.64	0.043
	[0.03]	[0.04]	[0.050]
Any spouse attended FFS	0.83	0.84	0.012
	[0.02]	[0.03]	[0.038]
Any spouse attended gender training	0.9	0.9	-0.004
	[0.02]	[0.03]	[0.031]
Any spouse attended VSLA	0.76	0.83	0.065
	[0.02]	[0.03]	[0.042]
Joint decision-making index	2.06	2.24	0.173*
	[0.06]	[0.07]	[0.099]
Labor index	0.91	0.95	0.040***
	[0.01]	[0.01]	[0.014]
Income control index	0.47	0.52	0.054*
	[0.02]	[0.03]	[0.032]
Distraction treatment	0.5	0.41	-0.093*
	[0.03]	[0.04]	[0.051]
Experiment order	10.27	9.4	-0.871
	[0.37]	[0.57]	[0.669]
N	310	135	445

Significance levels: * < 10% ** < 5% *** < 1%

Standard errors in brackets

Appendix B: Robustness Checks

Table A.B1: Conditional Differences in Means by Decision-Maker: Male Base

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Female Decision-Maker (= 1)	-0.101*	0.0429	-0.266**	-0.387*	-175.9*	-0.0467	-232.0
	(0.0607)	(0.114)	(0.131)	(0.229)	(90.89)	(0.0289)	(152.0)
Joint Decision-Maker (= 1)	-0.0133	0.160	-0.222**	-0.162	-66.23	-0.0365	-142.7
	(0.0545)	(0.103)	(0.111)	(0.194)	(82.85)	(0.0259)	(145.2)
Order Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses.
Male decision-makers are the base category.

(d) for discrete change of dummy variable from 0 to 1

="* p<0.10 ** p<0.05 *** p<0.01"

Table A.B2: Conditional Differences in Means by Decision-Maker: Female Base

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Male Decision-Maker (= 1)	0.101*	-0.0429	0.266**	0.387*	175.9*	0.0467	232.0
	(0.0607)	(0.114)	(0.131)	(0.229)	(90.89)	(0.0289)	(152.0)
Joint Decision-Maker (= 1)	0.0880*	0.117	0.0439	0.225	109.6	0.0102	89.32
	(0.0518)	(0.0985)	(0.123)	(0.219)	(74.40)	(0.0249)	(124.2)
Order Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses.
Female decision-makers are the base category.

(d) for discrete change of dummy variable from 0 to 1

="* p<0.10 ** p<0.05 *** p<0.01"

Table A.B3: Conditional Differences in Means by Laborer: Male Base

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Female Laborer (= 1)	0.0813*	0.0223	0.153	-0.569***	132.6*	0.0271	412.6***
	(0.0482)	(0.0932)	(0.109)	(0.183)	(71.39)	(0.0236)	(120.2)
Joint Choice of Laborer (= 1)	0.105*	0.153	0.0157	-0.226	136.8	0.00326	279.0**
	(0.0590)	(0.117)	(0.148)	(0.243)	(84.32)	(0.0305)	(141.2)
Order Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses.
Male laborers are the base category.

(d) for discrete change of dummy variable from 0 to 1

="* p<0.10 ** p<0.05 *** p<0.01"

Table A.B 4: Conditional Differences in Means by Laborer: Female Base

	(1) Total Cups	(2) Grey Cups	(3) Red Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Male Laborer (= 1)	-0.0813*	-0.0223	-0.153	0.569***	-132.6*	-0.0271	-412.6***
	(0.0482)	(0.0932)	(0.109)	(0.183)	(71.39)	(0.0236)	(120.2)
Joint Choice of Laborer (= 1)	0.0241	0.131	-0.137	0.343	4.283	-0.0239	-133.6
	(0.0570)	(0.108)	(0.143)	(0.247)	(85.47)	(0.0304)	(148.4)
Order Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	521	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

="* p<0.10 ** p<0.05 *** p<0.01"

Table A.B5: Conditional Differences in Means by Sub-Optimal Labor Allocation

	(1) Total Cups	(2) Red Cups	(3) Grey Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment

Sub-Optimal Laborer (= 1)	-0.108**	-0.0735	-0.122**	0.273**	-134.7*	-0.0134	-326.3**
	(0.0463)	(0.0704)	(0.0615)	(0.122)	(81.53)	(0.0264)	(134.8)
Order Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	445	445	445	445	445	445	445

Marginal effects; Standard errors in parentheses
Base is the optimal laborer being assigned to sort

(d) for discrete change of dummy variable from 0 to 1

="* p<0.10 ** p<0.05 *** p<0.01"

Table A.B6: Conditional Differences in Means by Sub-Optimal Labor Allocation and Decision-Maker

	(1) Total Cups	(2) Red Cups	(3) Grey Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Sub-Optimal Laborer (= 1)	-0.0238	-0.0600	0.0455	0.551***	-47.56	-0.0223	-459.5**
	(0.0765)	(0.113)	(0.105)	(0.188)	(140.7)	(0.0456)	(232.4)
Female Decision-Maker	-0.0879	-0.230***	0.0532	-0.206	-152.3	-0.0511	-369.4**
	(0.0600)	(0.0889)	(0.0823)	(0.189)	(107.2)	(0.0347)	(177.0)
Joint Decision-Maker	0.0298	-0.246***	0.281***	0.0201	-29.46	-0.0397	-143.1
	(0.0589)	(0.0878)	(0.0809)	(0.179)	(108.9)	(0.0353)	(179.8)
Sub-Optimal Laborer (= 1) # Female Decision-Maker	-0.0531	-0.121	-0.0356	-0.447	-78.08	0.0152	471.1
	(0.110)	(0.168)	(0.148)	(0.295)	(196.6)	(0.0637)	(324.6)
Sub-Optimal Laborer (= 1) # Joint Decision-Maker	-0.203*	0.0914	-0.452***	-0.472*	-176.3	0.0131	-50.32
	(0.111)	(0.165)	(0.150)	(0.283)	(196.7)	(0.0637)	(324.8)
Order Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	445	445	445	445	445	445	445

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

="* p<0.10 ** p<0.05 *** p<0.01"

Table A.B 7: Conditional Differences in Means by Sub-Optimal Labor Allocation and Laborer

	(1) Total Cups	(2) Red Cups	(3) Grey Cups	(4) Unsorted Cups	(5) Expected Return	(6) Investment Std. Dev.	(7) Payment
Sub-Optimal Laborer (= 1)	-0.124**	0.0285	-0.232***	-0.147	-115.8	-0.00303	-201.4
	(0.0625)	(0.0952)	(0.0830)	(0.161)	(109.9)	(0.0356)	(180.1)
Female Laborer (= 1)	0.00175	0.152*	-0.120	-0.684***	52.49	0.0200	281.0*
	(0.0551)	(0.0831)	(0.0739)	(0.162)	(99.10)	(0.0321)	(162.4)
Sub-Optimal Laborer (= 1) # Female Laborer (= 1)	0.223	-0.0618	0.451**	-0.694	202.0	0.0336	759.7*
	(0.154)	(0.225)	(0.211)	(0.624)	(280.2)	(0.0906)	(459.1)
Order Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	445	445	445	445	445	445	445

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

="* p<0.10 ** p<0.05 *** p<0.01"

Appendix C: The Labor Cost Treatment

Half of the sessions included music and snacks in the bean-sorting area. The intention was that these would provide a distraction or an incentive to not work and therefore, adjust the labor costs of the participants. However, the intervention did not have the intended effect and was too weak to effect any changes. We encourage further research along these lines in real-effort tasks to find ways to exogenously affect labor costs. This is particularly important in gender studies because women tend to have higher marginal labor costs because they have more household duties than men. Table A.C1 to A.C3 show that the labor cost treatment had no overall effects, no heterogeneous effects by decision-maker, and no heterogeneous effects by laborer.

Table A.C1: Effect of Labor Cost Treatment on Experiment Outcomes

	(1) Total Cups	(2) Red Cups	(3) Grey Cups	(4) Unsorted Cups	(5) Investment Std. Dev.	(6) Payment
Labor Cost Treatment	0.0597	-0.0648	0.153	-0.168	-0.00754	103.4
	(0.0948)	(0.158)	(0.124)	(0.206)	(0.0303)	(152.0)
Order Controls	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No
N	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * p < 0.10, ** p < 0.05, *** p < 0.01

Standard errors clustered at session-treatment group level

Table A.C2: Effect of Labor Cost Treatment on Experiment Outcomes by Decision-Makers

	(1) Total Cups	(2) Red Cups	(3) Grey Cups	(4) Unsorted Cups	(5) Investment Std. Dev.	(6) Payment
Labor Cost Treatment	0.0436	-0.237	0.309	-0.202	-0.0473	173.2
	(0.109)	(0.210)	(0.211)	(0.346)	(0.0502)	(243.8)
Woman	-0.108	-0.355**	0.132	-0.287	-0.0748*	-161.8
	(0.0888)	(0.179)	(0.153)	(0.245)	(0.0403)	(237.4)
Joint	-0.0231	-0.342**	0.269	-0.244	-0.0613	-72.38
	(0.0896)	(0.174)	(0.198)	(0.297)	(0.0452)	(202.5)
Labor Cost Treatment # Woman	0.0259	0.272	-0.214	-0.127	0.0715	-116.5
	(0.130)	(0.241)	(0.233)	(0.496)	(0.0530)	(285.0)
Labor Cost Treatment # Joint	0.0233	0.251	-0.208	0.156	0.0457	-80.28

	(1) Total Cups	(2) Red Cups	(3) Grey Cups	(4) Unsorted Cups	(5) Investment Std. Dev.	(6) Payment
	(0.134)	(0.242)	(0.248)	(0.449)	(0.0574)	(253.6)
Order Controls	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No
N	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1. * p <

0.10, ** p < 0.05, *** p < 0.01

Standard errors clustered at session-treatment group level

Table A.C3: Effect of Labor Cost Treatment on Experiment Outcomes by Laborers

	(1) Total Cups	(2) Red Cups	(3) Grey Cups	(4) Unsorted Cups	(5) Investment Std. Dev.	(6) Payment
Labor Cost Treatment	0.0101	-0.101	0.0936	-0.201	-0.00651	113.8
	(0.0976)	(0.227)	(0.172)	(0.239)	(0.0343)	(192.1)
Woman	0.0363	0.103	-0.0222	-0.618**	0.0177	416.4**
	(0.0702)	(0.176)	(0.119)	(0.273)	(0.0330)	(203.4)
Joint	0.0440	-0.116	0.160	-0.452	-0.0148	314.3
	(0.0791)	(0.199)	(0.176)	(0.278)	(0.0445)	(258.4)
Labor Cost Treatment # Woman	0.0758	0.0415	0.112	0.0408	0.000756	-6.285
	(0.0951)	(0.266)	(0.190)	(0.428)	(0.0480)	(252.6)
Labor Cost Treatment # Joint	0.0964	0.0703	0.0903	0.278	-0.0195	-142.5
	(0.120)	(0.342)	(0.242)	(0.522)	(0.0680)	(297.3)
Order Controls	No	No	No	No	No	No
Survey Controls	No	No	No	No	No	No
Session Fixed Effects	No	No	No	No	No	No
N	521	521	521	521	521	521

Marginal effects; Standard errors in parentheses. (d) for discrete change of dummy variable from 0 to 1.* p < 0.10,

** p < 0.05, *** p < 0.01

Standard errors clustered at session-treatment group level