The Effect of Gender and Gender Pairing on Bargaining: Evidence from an Artefactual Field Experiment

Ben D'Exelle, Christine Gutekunst, Arno Riedl

CES**ifo**

The Effect of Gender and Gender Pairing on Bargaining: Evidence from an Artefactual Field Experiment

Abstract

Men and women negotiate differently, which might create gender inequality in access to resources as well as efficiency losses due to disagreement. We study the role of gender and gender pairing in bilateral bargaining, using a lab-in-the-filed experiment in which pairs of participants bargain over the division of a fixed amount of resources. We vary the gender composition of the bargaining pairs as well as the disclosure of the participants' identities. We find gender differences in earnings, agreement and demands, but only when the identities are disclosed. Women in same-gender pairs obtain higher earnings than men and women in mixed-gender pairs. This is the result of the lower likelihood of disagreement among women-only pairs. Women leave more on the bargaining table, conditional on their beliefs, which contributes to the lower disagreement and higher earnings among women-only pairs.

Ben D'Exelle
University of East Anglia, School of Economics &
School of International Development
Norwich / Norfolk / United Kingdom
b.dexelle@uea.ac.uk

Christine Gutekunst
Maastricht University, Department of
Microeconomics and Public Economics
Maastricht / The Netherlands
christine.a.gutekunst@gmail.com

Arno Riedl*
Maastricht University, Department of
Microeconomics and Public Economics
Maastricht / The Netherlands
a.riedl@maastrichtuniversity.nl

*corresponding author

1. Introduction

Men and women tend to differ in their bargaining behavior. This has triggered a stream of studies on the link between gender and bargaining in different disciplines of the social sciences (for comprehensive literature reviews, see Kolb 2009, Hernandez-Arenaz and Iriberri 2019). Most evidence supports that women tend to be disadvantaged, because they are more reluctant to initiate a negotiation (see e.g. Babcock and Laschever 2003, Small et al. 2007, Bowles et al. 2007, and Kugler et al. 2018, for a meta-analysis of the psychological literature) or, because they are less successful in negotiations (see e.g. Kray and Thompson 2004, Keaveny and Inderrieden 2000, and the meta-analysis of Mazei et al. 2015, of the psychological literature).

An important area where the evidence is inconclusive, and thus more research is needed, is how gender differences in bargaining interact with gender pairing. Eriksson and Sandberg (2012) find that women are less likely to initiate a negotiation than men, but only when they are matched with a woman. Hernandez-Arenaz and Iriberri (2018) observe that women demand less in a bargaining game in the field, but only when matched with a man. Using an ultimatum game, Eckel and Grossman (2001) observe that men accept more offers from women than from men. Solnick (2001), in contrast, finds that both male and female respondents in an ultimatum game choose a higher minimum acceptable offer when matched with a female proposer instead of a male proposer.[1]

In this paper, we generate, first, new experimental evidence on the influence of gender and gender pairing on demands, agreements and earnings in bilateral bargaining, using a non-student participant sample. Second, we shed light on the mechanism behind bargaining differences that are induced by gender and gender pairing, by investigating how beliefs and the relation between beliefs and demands in bargaining are affected by gender and gender pairing.

We conducted an artefactual field experiment in rural Uganda. Participants were randomly paired and bargained over a given resource in a Nash demand game (NDG) (Nash 1953). Specifically, both participants in a pair simultaneously demanded a share of a given amount of resources. If the sum of demands did not exceed the amount of resources available they received their respective demands, otherwise they received a low disagreement payoff. This game

---

[1]In a bargaining-like environment, Castillo et al. (2013) observe in a field experiment, that male taxi drivers accept lower fares from female customers than from male customers.

combines simplicity with the possibility to observe bargaining behavior (i.e., the demands) and beliefs about the demands of the bargaining opponent as well as bargaining outcomes (i.e., agreed shares and disagreements).

To identify the role gender and gender pairing may play in bargaining we implemented two treatments where we varied the information about the bargaining opponent. In the baseline treatment, participants were not told whom they were paired with. This treatment allows us to test whether men and women bargain differently, exploiting the natural gender variation in the sample. In a second treatment, the identities of the participants were disclosed which allows us to test whether the gender of the opponent influences bargaining and whether this differs between men and women. Note that to avoid experimenter demand effects we did not actively tell participants the gender of the counterpart in this treatment. Rather, the disclosure of their general identities did give participants the information necessary to identify the gender of their opponent. Importantly, in our analysis we control for a large set of individual characteristics that are expected to differ between men and women, and at the same time may influence bargaining behavior.[2]

As bargaining behavior is likely affected by expectations about the opponent's behavior, we elicited individual beliefs about the demands of the opponent. This allows us to analyze how demands and bargaining outcomes depend on beliefs and whetherthe effect of beliefs differs between genders. In this analysis we use two concepts that link beliefs with demands and that could explain potential gender differences in bargaining behavior. First, men and women might differ in their ability to accurately predict their opponent's behavior. To capture this, we measure 'belief accuracy', which describes how close one's stated belief is to the actual demand made by the bargaining opponent. Second, we examine 'demand consistency', which captures the amount one leaves on the bargaining table conditional on one's belief.

Our results can be summarized as follows. First, we find gender differences in earnings, agreement and demands, but only when the identities of the paired participants are disclosed to one another. Second, when paired with a woman, women obtain higher earnings than men. In addition, women obtain higher earnings when paired with a woman than when paired with a man. These gender differences in earnings are the result of a lower likelihood of disagreement

---

[2]For instance, there is evidence that men and women differ in risk aversion (Eckel and Grossman 2008, Croson and Gneezy 2009, Charness and Gneezy 2012).

among women-only pairs than among mixed-gender pairs. Fourth, we do not find any gender differences in beliefs or belief accuracy. Men and women equally underestimate their opponent's demands. However, we do find that women's demands tend to be more inconsistent than men's: they leave more on the bargaining table, conditional on their beliefs. This counteracts the general underestimation of the opponent's demands and explains the relatively low disagreement rate among women-only pairs and the higher earnings obtained by women in such pairs.

Our study makes the following contributions to the literature. First, we provide new insights on the effect of gender and gender matching on bargaining outcomes. Specifically, the observed differences in earnings and disagreements suggest that women are not always worse off than men in bilateral bargaining settings, especially when the gender of the opponent is known. When they are matched with another woman, women might actually be better off than men. Second, our results on the role of beliefs and how they relate to individual demands, generate additional insights. In particular, women's smaller demand consistency counteracts the general underestimation of others' demands and leads to lower disagreement rates and higher earnings among women-only pairs. Third, most studies that analyze gender differences in bargaining behavior use WEIRD samples (Western, Educated, Industrialized, Rich and Democratic). It has been argued that such samples are unusual compared with the rest of the world population and that it is valuable to provide evidence from non-WEIRD samples (Henrich et al. 2010). We contribute to this endeavor by collecting data from a sample of men and women in rural Uganda.[3] Moreover, where gender differences in bargaining create gender inequality and efficiency losses, our insights could be of high relevance for development policy.[4]

The remainder of this paper is structured as follows. Section 2 introduces the design of the study, including a detailed description of the experiment and the socioeconomic survey, as well as their implementation. Section 3 describes the results, focusing on the influence of gender and gender pairings on earnings, disagreement rates, and demands. Section 4 analyzes the role of beliefs, and particularly of demand consistency and belief accuracy. Finally, Section 5 concludes.

---

[3]That there is substantial cross-cultural variation in gender differences in negotiation is confirmed by the meta-analysis of Shan et al. (2019).

[4]Gender inequality and poverty are particularly severe and intertwined in developing countries, such that tackling gender inequality has large potential to reduce poverty (World Bank 2012).

2. Study Design

In this section, we first describe the bargaining game we conducted, the treatment variations, and the socioeconomic survey. Thereafter, we explain the implementation of the study, including the study location, the sampling process, and the experimental procedures.

2.1. The Bargaining Game

In the experiment we implemented the well-known Nash demand game (NDG) in which participants only have to make one decision, namely to state their demands over a given resource (Nash 1953). We chose this game because it combines simplicity with the possibility to observe bargaining behavior (i.e., the demands) as well as bargaining outcomes (i.e., agreed shares and disagreements). As we conducted the experiment with a non-student sample (see Section 2.4 for details) it was important to guarantee easy understanding of the rules of the game. In addition, in the NDG it is straightforward to elicit beliefs about the opponent's bargaining behavior because the opponent is in exactly the same strategic situation as the participant stating the beliefs.

In the game, two players i and j simultaneously and independently make demands x_i and x_j regarding a given resource R. If the sum of the demands made by the two players does not exceed the amount of the available resource, that is, $x_i + x_j \leq R$, then they reach an agreement and the players receives their respective demand. Otherwise, each player gets ex ante defined and known disagreement earnings $d_i = d_j = d$.[5] Let $u_i, u_j : [0, R] \cup \{d\} \rightarrow \mathbb{R}$ be players' von Neumann-Morgenstern utility functions which map from the space of possible demands into the payoff space. The set of possible bargaining agreements S is given by those (x_i, x_j) that fulfill $u_i(x_i) \geq u_i(d), u_j(x_j) \geq u_j(d)$ and $x_i + x_j \leq R$. It is easy to see that any pair of demands $(x_i, x_j) \in S$ which satisfy $x_i + x_j = R$ constitutes a Nash equilibrium.

These Nash equilibria are all efficient. However, as Harsanyi (1962) already pointed out, in real bargaining efficiency is not guaranteed because in equilibrium demands and beliefs need to be mutually consistent. Given the high strategic uncertainty, this is not straightforward and disagreements or inefficient agreements can be expected.

---

[5]Generally, $d_i \neq d_j$, however, in our experiment both players have symmetric disagreement earnings.

2.2. Treatments — Information Disclosure

In the experiment participants were randomly matched into bargaining pairs. In a bargaining pair a woman could be matched with another woman (women - women pair, WW in short), a man could be matched with another man (men - men pair, MM), and each gender could be matched with the opposite gender in a mixed-gender pair. In the analysis we distinguish between mixed women - men pairs (WM) and mixed men - women pairs (MW) depending on whose perspective we take when describing behavior. That is, in a WM pair the woman is the so-called 'ego' while the man takes the role of 'alter', and vice versa for MW pairs.

To understand the influence of gender and gender pairing on bargaining behavior and outcomes, our experimental design combines the natural variation in participants' gender with the experimental variation of gender pairing and information disclosure about the identity of the bargaining opponent. In a baseline anonymity treatment (AN) no information about bargaining partners was revealed. Participants, thus, did not know the gender of their bargaining opponent. In contrast, in the full disclosure treatment (FD) each participant in a bargaining pair was informed about the other's identity and, thus, also gender.[6] Treatment AN allows us to investigate the effect of own gender in bargaining when not knowing the gender of the bargaining opponent, and with treatment FD we can analyze the potential effect of gender pairing next to an own gender effect.

In FD, we disclose the participants' identities by showing the picture and name of both players in a pair, but without actively mentioning their gender. An alternative approach to identify an effect of gender pairing would be to actively tell participants the gender of the counterpart, thereby risking the introduction of an experimenter demand bias. Our approach has the advantage of avoiding such a bias, but bears the disadvantage that a number of other characteristics are disclosed in addition to gender. Therefore, in the analysis we control for potential confounding factors with the help of detailed information from a socioeconomic survey.

---

[6]In a separate treatment we implemented a semi-disclosure condition where only one participant in a pair has information about the other's identity. This treatment was conducted with a different group of participants and is not considered in this paper.

2.3. Socioeconomic Survey

A few weeks before the experiment, we interviewed each participant in private to capture important socioeconomic and psychological characteristics of participants and their households. During the interview, we asked questions on wealth, age, education, gender, and social ties. To measure household wealth, we asked a variety of questions on the characteristics of the home they lived in (number of rooms, type of flooring, etc.), their access to electricity, and how much livestock the household owned. Age and education were elicited as continuous variables, with education being measured in years of schooling. To elicit social ties, participants were shown individual photographs of each of the other participants from their village and asked whether they were friends with the respective person.[7]

Further, we asked a variety of questions to elicit trust, agreeableness and risk aversion. Regarding trust, we asked the respondents whether they thought that most people could be trusted, whether most people would try to take advantage if they had the chance to, and if most of the time people tried to be helpful (Dohmen et al. 2008, 2012, Altmann et al. 2008). All three questions had three answer options, equivalent to 'agreeing fully,' 'agreeing somewhat,' and 'disagreeing fully.'[8] Risk attitudes were captured through a self-reported score, detailing whether respondents 'Take risk a lot', 'Take risk but not a lot', 'Avoid risk but not a lot', or 'Avoid risk a lot' (Dohmen et al. 2005, 2012). Agreeableness is based on the questions from the Big Five questionnaire (Costa and McCrae 1992). Specifically, respondents were asked whether they agreed with the following statements: 'I try to forgive and forget when insulted'; 'I am ready to fight back if somebody else starts a fight'; 'I hesitate to express anger even if it is justified'. Answers were given on a five-point scale ranging from strongly disagree to strongly agree.

---

[7]The main advantage of social-ties elicitation via a list that contains each of the other participants from the village is that a reporting bias due to the size of a respondent's network can be avoided. More connected people might otherwise be more likely to forget a social tie, but by showing pictures each participant gets the same cues to allow him/her to remember all his/her social ties.

[8]For the question 'Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?' answer options were 'Most people can be trusted'. 'Some can be trusted, some not', and 'Can't be too careful'. For the question 'Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?' answer options were 'Most try to take advantage', 'Some take advantage, some not', and 'Most are fair'. For the question 'Would you say that most of the time people try to be helpful, or that they are just looking out for themselves?' answer options were 'Most are helpful', 'Some are helpful, some are not', and 'Most look after themselves'.

2.4. Implementation of the Survey and Experiment

The study was conducted in the Sironko district in eastern Uganda. In this district, a sample of participants was chosen, using a multicluster sampling approach. In the first step, we selected a random set of 24 villages. The selected villages were small-scale, close-knit communities of the type that is common in the rural areas of most Sub-Saharan African countries. In each of the villages, we selected a random sample of households and from each of these selected households, one adult was chosen at random to participate in the survey and experiment.

The household survey was conducted in private at the participants' houses, using portable data entry devices. The experiment was organized in sessions for which participants from at most two villages where gathered in one common place. Individual decision cards were used to record participants' demands and their beliefs about the demand of their counterpart. In AN, no other information was given or entered in the decision card. In contrast, in FD, to disclose identities, a decision card displayed the participant's own name and picture as well as the counterpart's name and picture, to show that both participants had information about each other. Each participant took two decisions in the AN condition and two decisions with different opponents in the FD condition.[9] This allows us to observe variation in behavior of one participant towards different opponents of possibly differing gender. Participants were randomly re-matched into new pairs for each decision, and a new decision card was distributed for each of these decisions. To avoid learning effects, we gave no feedback between decisions. Matching of pairs, as well as preparation of the decision cards was performed before the experiment started.

Beliefs and demands were elicited through two simple questions on the decision cards that asked participants for the amount that they expected the other one to demand, and the amount that they demanded for themselves, respectively. The resource available was 16000 Ugandan Shilling (UGX), which was roughly equal to two daily wages for the average participant. The disagreement payoff was 2000 UGX. Demands and beliefs could be stated freely as integers between 0 and 16000. Demands were incentivized by informing the subjects that, at the end of the experiment, one of their decisions would be selected at random to be paid out. This

---

[9]In FD both opponents were from the same village. In AN participants were paired with someone who lived in the same village for one decision and with an opponent who lived in a different village for the other decision. This was implemented to see if it matters whether the opponent lives in the same or a different village. For this paper, we only use the same village pairs.

protocol of randomly selecting one single decision for payment was followed to avoid hedging between different decisions (for theoretical arguments and empirical evidence in favor of this procedure, see Azrieli et al. 2018, 2020). Belief elicitation was not incentivized.[10]

Instructions were read out loud and control questions on how to calculate own and others' earnings were asked in private (for details see Appendix D). If participants were unable to answer the control questions further clarifications were given in private. If they were still unable to answer the control questions, a note was made such that we could exclude their decisions from the analysis. About 8% of our participants had problems understanding the instructions and were excluded from the analysis. Participants were seated at desks with sufficient distance from each other, so that they were unable to read their neighbors' decision cards. They were instructed not to talk to their neighbors and to keep their decisions confidential during the experiment. Participation was entirely voluntary, and participants were told that they were free to leave the experiment at any point. Nobody made use of this option. At the end of the experiment, payments were received in private. Payments received as well as any decisions made by the participants were treated confidentially, and this was clearly explained at the beginning of the experiment.

Each session took about two hours. 342 decisions were made overall, out of which 180 in AN and 162 in FD. In total, 180 participants took part in the experiment, out of which 81 participated in FD.[11] Of those individual decisions, in AN 97 decisions were taken by a woman and 83 decisions were taken by a man; in FD 78 decisions were taken by a woman and 84 decisions were taken by a man. The slight imbalance in male and female participants is a result of the random selection at household level. In total, in AN there were 26 WW pairs, 19 MM pairs, and 45 mixed pairs, and in FD there were 23 WW pairs, 26 MM pairs, and 32 mixed pairs.

---

[10]We chose not to incentivize beliefs because there is evidence that incentivized belief elicitation entails the risk that participants hedge between action and beliefs, especially when they have a financial stake in the predicted action, which is the case in our set-up (Rutström and Wilcox 2009, Blanco et al. 2010, Armantier and Treich 2013).

[11]This number is conditional on both participants in a pair having fully understood the instructions. This excludes about 8% of all participants, and their respective partners in a pair, because of difficulties understanding the instructions. Numbers in AN include participants that would go on to take part in the separate semi-disclosure condition as mentioned in Footnote 6.

3. Results

In the following, we describe the role of gender and gender pairing in determining individual earnings, pairwise disagreement outcomes, and individual demands. Subsequently, we discuss the role of beliefs, and particularly the relation between beliefs and demands, as a possible channel for observed gender differences.

3.1. Earnings

Figure 1 displays individual earnings by information condition, gender, and gender pairing. Earnings of women are shown in the upper row, earnings of men in the bottom row. The panels in the left column display earnings in AN separated by gender. The middle and right columns display earnings in FD. For FD, where bargaining opponents knew each other's gender, individual earnings are displayed by gender pairings. The middle column displays men's and women's earnings when in a mixed-gender pair and the right column shows men's and women's earnings when in a same-gender pair.
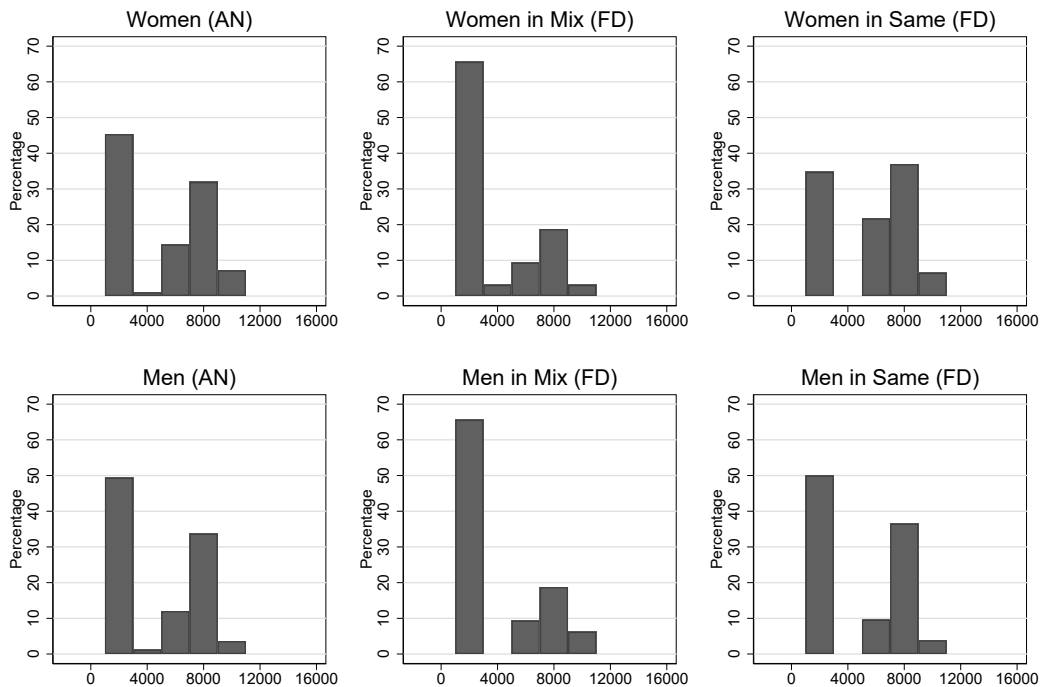
N= 342 (162 FD, 180 AN). Histograms show distribution of earnings by treatment and gender/ gender pairing; left-most bars also indicate frequency of disagreements.

Figure 1: Distribution of Individual Earnings

9

In each panel, the left-most bar reflects the frequency of disagreement with a payoff of 2000 UGX. The other bars display individual earnings in case of agreement. In the following we refer to earnings including disagreements as 'final earnings' and to earnings from agreements only as 'agreed earnings'. Visual inspection of the figure suggests very similar earnings for men and women in AN. In contrast, in FD some interesting differences are discernable. Perhaps most eye-catching is the relatively low rate of disagreement earnings in women–women pairs and the relatively high rate of disagreement earnings in mixed gender pairs. This translates into some substantial differences in 'final earnings' between genders across gender pairings. For instance, on average women earn 5543 UGX in same gender pairs, whereas men earn only 3813 UGX in mixed-gender pairs. In contrast, 'agreed earnings' appear to be similar between genders. Indeed, across all gender pairings agreed earnings vary only between 7273 UGX and 7636 UGX (for detailed descriptive statistics of individual earnings, see Table B.1).

To test whether there are any statistically significant differences between genders, we use the following regression specification separately for AN and FD:

$$y_{ij} = \beta_0 + \beta_1\, G_{ij} + \beta_2\, X_{ij} + \mu_k + \varepsilon_{ij}, \tag{1}$$

where y_{ij} are individual earnings of ego i when matched with alter j. In the AN regressions, G_{ij} is a dummy variable that is equal to one when ego is a man, zero otherwise. In the FD regressions, G_{ij} is a vector of dummy variables capturing the different gender pairings WW, MM, MW and WM. Note that an ego-man can be in a same-gender pair (MM) or in a mixed-gender pair (MW) and an ego-woman can be in a same-gender pair (WW) or in a mixed-gender pair (WM). X_{ij} represents a vector of socioeconomic controls, for FD related to ego's and alter's characteristics and for AN only related to ego's characteristics.[12] This vector includes wealth, age, education, trust, and risk aversion of ego (and alter for FD). The variables for wealth and trust are indices that are obtained from a principal component analysis (see Appendix A.1 for details). Age and education are measured in years, and risk aversion is measured on a four-point scale with a larger value reflecting higher risk aversion. We only include controls for which gender differences were observed.[13] μ_k stands for experimenter fixed effects, and ε_{ij}

---

[12]In the AN regressions regarding ego's individual earnings, we do not include alter's gender and other characteristics because they are unknown to ego by definition.

[13]See Appendix Table A.1 for a detailed comparison between men and women of the socioeconomic characteristics as elicited during the household survey. No gender differences are found in agreeableness and network ties.

captures the remaining idiosyncratic error. We use bootstrapped cluster-robust standard errors with 2000 repetitions to adjust for potential non-independencies within sessions.[14]

Table 1 reports the regression results, for AN in the upper panel and for FD in the lower panel. Results for final earnings are shown in Models 1 and 2 and for agreed earnings in Models 3 and 4, using the sub-sample of pairs of participants where an agreement was reached. Models 1 and 3 provide results without controls, Models 2 and 4 include the mentioned socioeconomic controls for ego (and for alter in FD). For AN, women form the reference category against which earnings of men are compared. For FD, the reference category is WW.

For AN, the regression results show no significant gender differences either in final earnings or in agreed earnings, indicating that own gender does not affect earnings when the gender of the bargaining counterpart is unknown. In FD results are different, however. In Models 1 and 2, the significant and negative coefficient of WM shows that women in a mixed-gender pair have substantially lower final earnings than women in a same-gender pair (1788.8 UGX, Model (2)). Further, the significant and negative coefficient of MW shows that men who are paired with a woman earn less than women who are paired with a woman (1499.7 UGX, Model (2)). On the other hand, for men, the gender of the counterpart does not influence final earnings (none of the MM vs. MW and MM vs. WM comparisons are statistically significant, as reported in the panel below the regression table). Finally, Models 3 and 4 show that agreed earnings do not depend on the gender of the counterpart for either men or women.

We summarize the discussed findings in our first result.

Result 1. *(A) In AN, neither final nor agreed earnings differ between genders. (B) In FD, women's final earnings are lower when paired with a man than when paired with a woman. Men's final earnings do not depend on the gender of their counterpart. However, when paired with a woman, men earn less than women. (C) No gender differences are observed for agreed earnings in FD.*

This result indicates that, when identities are revealed, either side of a mixed-gender pair fares worse than same-gender female pairs. Put differently, when knowing that they are matched

---

[14]Here we follow Cameron et al. (2008) who advise to use bootstrapping if the number of clusters is low. We ran 11 sessions which consitute the amount of clusters in the regression.

Table 1: The Effect of Gender (Pairings) on Individual Earnings

	Final Earnings		Agreed Earnings	
	AN			
	(1)	(2)	(3)	(4)
Men	-424.6	-522.4	-296.7	-501.7
	(480.6)	(519.9)	(369.9)	(344.2)
Constant	5889.4***	6701.2***	7735.1***	8233.9***
	(397.2)	(1250.2)	(231.5)	(822.8)
Observations	180	179	96	95
R^2	0.063	0.079	0.014	0.181
	FD			
	(1)	(2)	(3)	(4)
MM	-751.0	-244.2	149.7	31.18
	(764.4)	(1135.7)	(140.2)	(297.4)
WM	-1728.4***	-1788.8***	-161.0	-370.4
	(647.0)	(655.5)	(551.3)	(553.6)
MW	-1603.4***	-1499.7**	202.6	472.2
	(614.9)	(589.0)	(333.8)	(368.2)
Constant	5527.9***	3633.8*	7414.5***	6026.8***
	(656.5)	(2143.5)	(167.6)	(1024.4)
Observations	162	162	78	78
R^2	0.056	0.107	0.008	0.110
	Wald tests (p-values)			
MM - WM	0.340	0.200	0.563	0.409
MM - MW	0.396	0.299	0.887	0.359
Controls	No	Yes	No	Yes

Linear regressions. Dependent variables are final earnings (models 1 & 2) and earnings in the case of agreement (models 3 & 4). Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions) clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels. Controls include wealth, age, education, trust, and risk aversion of ego (and alter for FD).

with a woman, participants in WW pairs seem to be able to secure higher final earnings than those in mixed-gender pairs. In addition, the absence of gender differences in agreed earnings indicates that the observed gender differences in final earnings for pairs in FD are likely related to differences in disagreement rates between gender pairs. Next we test this conjecture.

3.2. Disagreement

We test for differences in disagreement rates between different gender pairings using a regression approach equivalent to the one laid out in Equation (1), again separately for AN and FD.[15] However, given that disagreement is an outcome on the pair level the mixed gender pairings MW and WM are pooled and the vector G_{ij}, thus, now includes the categories MM, WW, and mixed-gender pairs, for both AN and FD. In the regressions, mixed-gender pairs are the reference category, against which outcomes of WW and MM pairs are compared.

Table 2 provides the results of the regression analysis by gender pairing in AN and FD. Model 1 provides results without controls, while Model 2 includes ego and alter controls for wealth, age, education, trust, and risk aversion. In AN, no difference in disagreement between different gender pairings is observed. For FD, however, a gender pairing effect can be observed. The significant and negative coefficient of WW in both models shows that the probability of disagreement is more than 30 percentage points lower among WW pairs than among mixed-gender pairs. The coefficient of MM is not significant, which indicates that the likelihood of disagreement is not statistically different between MM pairs and mixed-gender pairs. We summarize these observations in our next result.

Result 2. *(A) In AN, no gender pairing differences in disagreements are observed. (B) In FD, the probability of disagreement is more than 30 percentage points lower among WW pairs than among mixed-gender pairs, while disagreement is equally likely in MM pairs and mixed-gender pairs.*

The lower disagreement rate of female same-gender pairs in FD resonates with women's higher final earnings, indicating that it is this lower disagreement rate that enables women in same-gender pairs to secure higher earnings than those in mixed pairs.

3.3. Demands

Given that the occurrence of disagreement depends on individual demands, similar gender differences can be expected for individual demands. Figure 2 provides histograms that display

---

[15]That is, here we apply a linear probability regression model. See Appendix Tables C.5 and C.6 for results based on a probit regression. The results are qualitatively the same.

Table 2: The Effect of Gender Pairings on Pairwise Disagreement

| | AN | |
	(1)	(2)
WW	-0.117	-0.168
	(0.119)	(0.128)
MM	-0.0268	-0.140
	(0.152)	(0.181)
Constant	0.372***	-0.0895
	(0.0881)	(0.613)
Observations	90	89
R^2	0.069	0.165
	Wald tests (p-values)	
WW - MM	0.608	0.895
	FD	
	(1)	(2)
WW	-0.308***	-0.347***
	(0.0998)	(0.116)
MM	-0.156	-0.263
	(0.182)	(0.236)
Constant	0.658***	1.093***
	(0.112)	(0.406)
Observations	81	81
R^2	0.064	0.163
	Wald tests (p-values)	
WW - MM	0.278	0.702
Controls	No	Yes

Linear regressions. Dependent variable is disagreement. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels. Controls include wealth, age, education, trust, and risk aversion of ego and alter.

the distribution of demands by information condition and gender. For AN, the left panels of the figure do not indicate a difference in demand behavior between genders. In FD, however, interesting differences can be detected. In particular, the mode for men in a mixed-gender pair (lower middle panels) is at 10000 UGX, whereas it is at the equal split of 8000 UGX in all other FD panels. This suggests that men's demand behavior in a mixed-gender pair is more aggressive than in other gender pairings.

To test whether differences are statistically significant, we run regressions with demands as dependent variable, using a regression model equivalent to the one described in Equation (1).
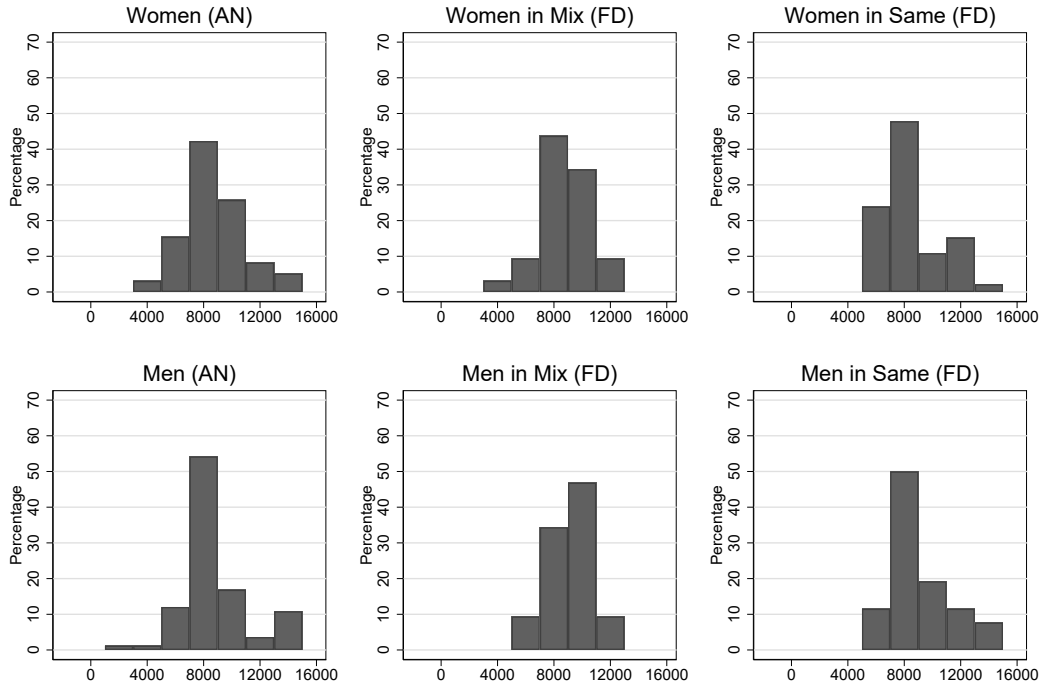
N= 342 (162 FD, 180 AN). Histograms show distribution of demands by treatment and gender/ gender pairing.

Figure 2: Distribution of Demands

Table 3 reports the results, for AN in the upper panel and for FD in the lower panel. Model 1 shows results without socioeconomic controls, whereas Model 2 includes wealth, age, education, trust, and risk aversion of ego (and of alter in FD). As in the earnings regressions, women form the reference category in the AN regressions, while WW is the reference category in the FD regressions.

In AN, no significant gender differences show, indicating that gender does not matter for demands if the identity of the counterpart is not known. In FD, however, important differences can be observed. The significant and positive coefficient of MW in both models shown in the lower panel indicates that men who are paired with a woman demand more than a woman who is paired with another woman. The insignificant coefficient of WM, however, shows that women do not demand significantly less in a mixed-gender pair than in a same-gender pair. Further, the coefficients of MM and WM are not statistically different from one another, which indicates that, if paired with a man, men and women make similar demands, on average (see the MM vs. WM comparison on the bottom of the FD panel). We summarize these observations in the following result.

Result 3. *(A) In AN demand behavior of men and women does not differ. (B) In FD, men demand more than women, but only when they are paired with a woman.*

This result is consistent with the observed gender differences in disagreement rates in FD. If men who know that they are in a mixed pair demand more, while women behave the same in a mixed and in a same-gender pair, more disagreements are likely to occur in mixed pairs. Men's higher demand in a mixed pair causes disagreement which leaves not only themselves, but also those women they are matched with, worse off.

Table 3: The Effect of Gender (Pairings) on Demands

	AN	
	(1)	(2)
Men	127.1	-74.73
	(320.1)	(322.4)
Constant	8223.0***	9187.3***
	(345.8)	(936.3)
Observations	180	179
R^2	0.021	0.079
	FD	
	(1)	(2)
MM	707.2	697.4
	(438.7)	(620.0)
WM	330.1	335.0
	(350.9)	(321.2)
MW	736.4***	844.5**
	(269.2)	(394.8)
Constant	7983.1***	7060.2***
	(364.6)	(1486.7)
Observations	162	162
R^2	0.028	0.052
	Wald tests (p-values)	
MM - WM	0.537	0.603
MM - MW	0.958	0.834
Controls	No	Yes

Linear regressions. Dependent variable is ego's demand. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels. Controls include wealth, age, education, trust, and risk aversion of ego (and alter in FD).

4.   The Role of Beliefs

To explore potential mechanisms behind the observed gender differences in earnings, agreements, and demands, we look into the role of beliefs. As Harsanyi (1962) points out, the NDG is a game that contains a large amount of strategic uncertainty with regard to the other's demand. Therefore, bargainers' beliefs about each other's demand are likely important determinants of their bargaining behavior and resultant outcomes.

Figure 3 displays the distribution of beliefs by information treatment and gender pairing. While the mode is 8000 for men and women in AN, as well as for same-gender pairs and for women in a mixed-gender pair in FD, the mode for men in a mixed-gender pair is 6000. This is in line with the observed demand mode of men in a mixed pair at 10000. Apparently, men expect the other to demand less when the other is known to be a woman, and they make their demands accordingly. Women, on the contrary, seem to expect the other to demand the same when their identity is revealed, irrespective of the opponent's gender. None of these visually observed differences in beliefs between gender in AN or between gender pairs in FD is statistically significant though (see Appendix Tables C.9 for AN and C.10 for FD).

The observation that beliefs do not differ across gender (and gender pairings in FD), while demands do, points to the possibility that gender differences in demands and earnings could be explained by potential gender differences in the relation between beliefs and demands. There are two concepts that link beliefs and demands: (1) 'belief accuracy', meaning that expected demands of the other match with their actual demands, (2) 'demand consistency', meaning that own demands and own beliefs are consistent with each other. The concept of belief (in)accuracy is prevalent in the existing experimental literature[16], while the second aspect, demand (in)consistency, has not been explored. In the following we take a closer look at belief accuracy and demand consistency, and whether they differ between genders.

Belief accuracy describes how close a belief b_{ij} of player i is to the actual demand x_j made by the opponent. A belief is accurate if it predicts the correct amount, in which case own beliefs and others' demands exhibit a correlation of 1. Demand consistency describes whether, given

---

[16]See for instance Gächter and Renner (2010), Fischbacher and Gachter (2010), and Croson (2007), as well as the broader literature on belief incentivization that routinely rewards participants for the accuracy of their beliefs.
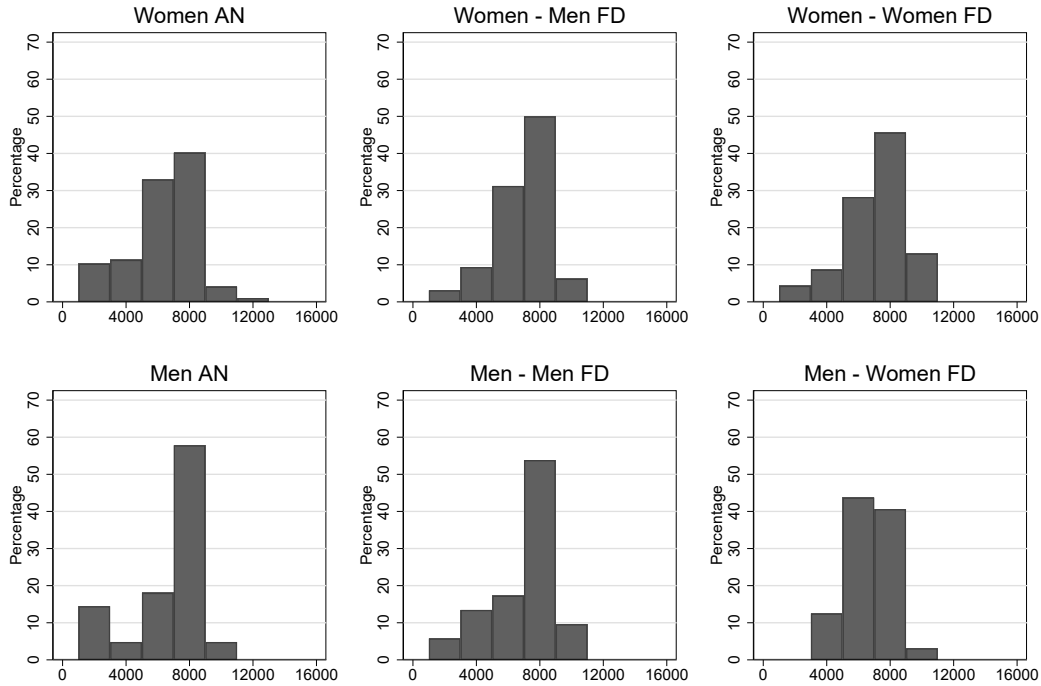
N= 342 (162 FD, 180 AN). Histograms show distribution of beliefs by treatment and gender/ gender pairing.

Figure 3: Distribution of Beliefs

a participant's belief, their own demand uses up all of the available resource. When demands are consistent, own beliefs and own demands exhibit a correlation of -1.[17]

To analyze whether there are gender differences in belief accuracy and demand consistency, Figure 4 displays scatterplots of beliefs and demands. These allow us to inspect whether any differences in belief accuracy and demand consistency are discernible. The left panels show own beliefs on the x-axes and other's demand on the y-axes, thus relating to belief accuracy, whereas the right panels show own beliefs on the x-axes and own demand on the y-axes, thus relating to demand consistency. The panels in the first row depict data for AN, with dark dots representing men and light dots representing women. The second and third row show behavior in FD, with panels in the second row representing women and panels in the third row

---

[17]In the neuroscience literature on Theory of Mind there is some evidence suggesting that women have better social cognition, which might give them an advantage in forming correct beliefs (see, for instance, McClure 2000, Baron-Cohen et al. 1999, Brackett and Salovey 2006, Adenzato et al. 2017). On the other hand, numerous studies have shown that women are, on average, more risk averse than men (see, for instance, Charness and Gneezy 2012, Croson and Gneezy 2009, Eckel and Grossman 2008). A stronger desire to avoid risks might increase the inconsistency of women's demands.

representing men. In the FD panels, the light dots indicate that alter is of the same gender, whereas the dark dots indicate that alter is of a different gender.
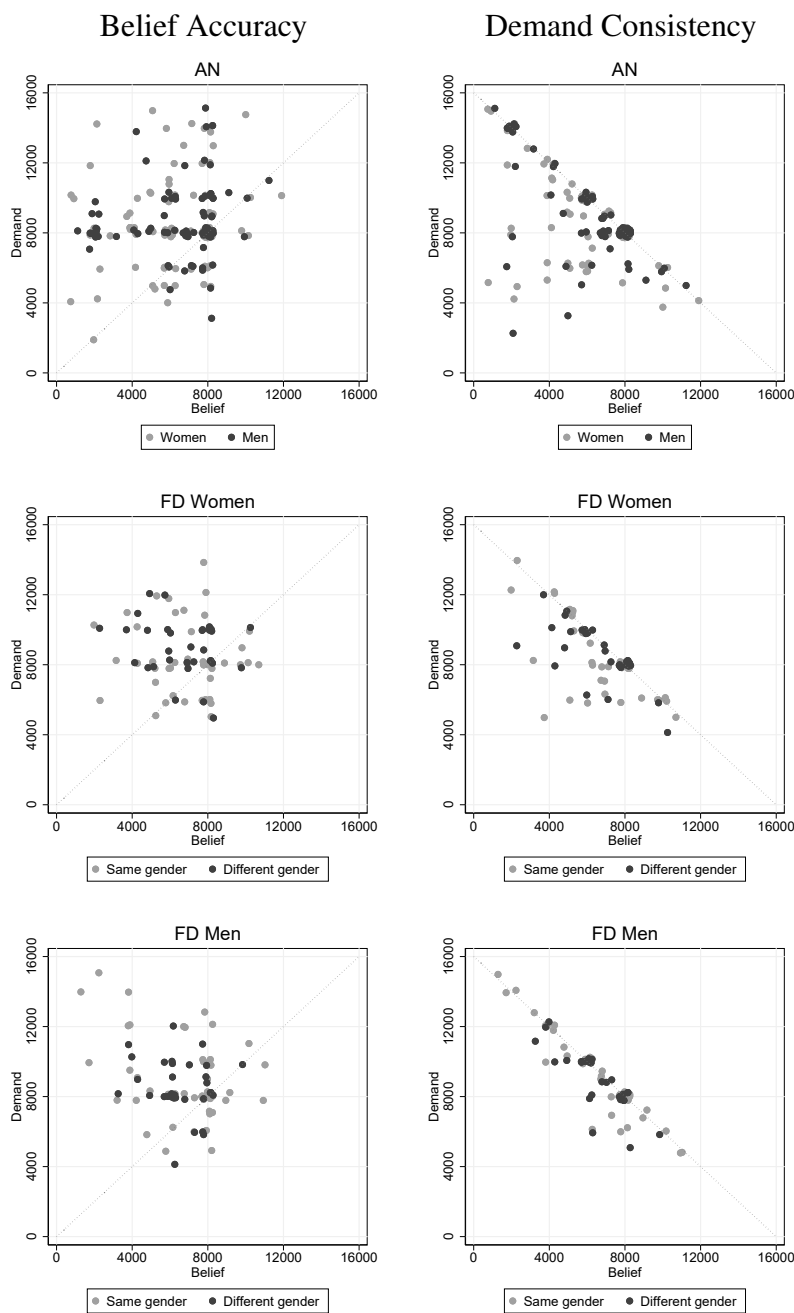
Figure 4: Scatter Plots for Belief Accuracy and Demand Consistency

For belief accuracy, the 45-degree line indicates where alter's demand and ego's belief are identical, that is, where beliefs are perfectly accurate. It can be seen that in AN and FD, for men as well as for women, there are many points to the left of the 45-degree line, and a few to the right. This indicates that participants generally tend to underestimate the opponent's demand. No difference between genders is discernible from the graphs either in AN or in FD.

For demand consistency, the negative 45-degree line indicates where ego's demand and ego's belief add exactly to 16000 UGX, that is, where demands are perfectly consistent. Also in this graph, many dots can be found to the left of the negative 45-degree line, indicating that demand and belief sum up to less than the available resource. For AN, no difference is observable between the panels for men and women. For FD, on the contrary, more points in the men's panel seem to be located on the negative 45-degree line than in the women's panel, suggesting that women's demands are less consistent than those of men.

To investigate whether any of the visually observed gender differences regarding belief accuracy or demand consistency are significant, we use a regression that extends Equation (1) by including ego's belief as an independent variable while maintaining demands as dependent variable. In regressing demands against beliefs, we directly follow the outline of Figure 4. Beliefs are interacted with the gender pairings to allow for a separate analysis of the role of beliefs for each of the possible gender pairings. For belief accuracy, the dependent variable is alter's demand. For demand inconsistency, the dependent variable is ego's demand.

Table 4 displays results for belief accuracy in Models 1 and 2, and for demand inconsistency in Models 3 and 4. As before, Models 1 and 3 show results without controls, while Models 2 and 4 display results with controls for ego's characteristics (and for alter's characteristics in FD). The upper panel displays results for AN, while the lower panel displays results for FD. In AN, men constitute the reference category, against which women's behavior is compared. In FD, MM is the reference category, against which behavior in the other gender pairings is compared.

As expected from Figure 4, no gender differences are found for AN in the upper panel for either belief accuracy or demand consistency. For FD, on the other hand, differences can be observed, particularly for demand consistency. For belief accuracy in Models 1 and 2 of the lower panel, the coefficient of Belief is not significant and there are no differences between gender pairings. This confirms the expectation that belief accuracy does not differ between gender pairings in FD. For demand consistency in Models 3 and 4, however, the coefficient of Belief is indistinguishable from -1 ($p = 0.692$, Wald test). Moreover, the constant term is indistinguishable from 16000 ($p = 0.811$, Wald test). Together, these findings indicate that for MM pairs demands are close to being consistent. For the gender combinations WW and WM both the constant terms and the belief coefficients are significantly different. The difference

Table 4: The Effect of Gender (Pairings) on Belief Accuracy and Demand Consistency

	Belief Accuracy		Demand Consistency	
	AN			
	(1)	(2)	(3)	(4)
Women	-713.1	-586.6	-1378.5	-1357.7
	(694.5)	(776.8)	(1583.2)	(1418.8)
Belief	0.123	0.122	-0.664***	-0.692***
	(0.0864)	(0.0963)	(0.173)	(0.169)
Women × Belief	0.104	0.0937	0.182	0.220
	(0.125)	(0.138)	(0.210)	(0.189)
Constant	7440.4***	6408.9***	13001.1***	13592.8***
	(626.8)	(1185.1)	(1331.8)	(1570.0)
Observations	180	179	180	179
R^2	0.051	0.070	0.312	0.359
	FD			
	(1)	(2)	(3)	(4)
WW	-2746.2	-2490.4	-2675.9**	-2578.6**
	(1981.7)	(2123.5)	(1185.8)	(1136.1)
WM	-943.7	-434.4	-3101.6*	-3270.6**
	(1890.9)	(1964.2)	(1640.3)	(1494.8)
MW	-1952.3	-1274.1	-1575.4	-1292.0
	(2150.2)	(2372.6)	(1204.8)	(1152.3)
Belief	-0.342	-0.328	-1.004***	-1.036***
	(0.234)	(0.256)	(0.113)	(0.0904)
WW × Belief	0.306	0.274	0.310**	0.321**
	(0.249)	(0.270)	(0.151)	(0.141)
WM × Belief	0.141	0.0881	0.395*	0.450**
	(0.235)	(0.250)	(0.223)	(0.207)
MW × Belief	0.229	0.127	0.209	0.188
	(0.282)	(0.305)	(0.158)	(0.147)
Constant	11158.2***	10288.9***	15937.2***	15740.4***
	(1987.2)	(2143.8)	(893.9)	(1087.3)
Observations	162	162	162	162
R^2	0.079	0.103	0.676	0.688
	Wald tests (p-values)			
WW - WM	0.144	0.185	0.840	0.718
WW - MW	0.512	0.325	0.422	0.317
WW×Belief - WM×Belief	0.360	0.386	0.766	0.619
WW×Belief - MW×Belief	0.663	0.419	0.575	0.421
Controls	No	Yes	No	Yes

Linear regressions. In Model 1 & 2, dependent variable is alter's demand. In Model 3 & 4, dependent variable is ego's demand. Standard errors in parentheses, clustered at session level, bootstrapped (with 2000 repetitions) at the session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels. Controls include wealth, age, education, trust, and risk aversion of ego (and alter in FD).

in the constant terms shows in the negative and (marginally) significant coefficients of WW and WM, while the difference in the belief coefficient shows in the positive and (marginally) significant coefficient of WW×Belief and WM×Belief. For both WW and WM, the belief coefficient (Belief + WW×Belief or Belief + WM×Belief) is clearly smaller in magnitude than -1 ($p = 0.025$ for WW and $p = 0.024$ for WM, Wald test). The constant terms (Constant + WW or Constant + WM), on the other hand, are clearly distinguishable from 16000 ($p = 0.039$ for WW and $p = 0.009$ for WM, Wald test).

Taken together, the analysis of the gender-specific constant terms and belief coefficients provides two important insights. First, on average, in FD demand inconsistency of women is higher than for men who are paired with a man. This holds irrespective of the gender of the woman's counterpart, as can be seen from the insignificant differences between WW and WM, as well as WW×Belief and WM×Belief, at the bottom of the FD panel. Second, the coefficients of WW and WW×Belief, as well as WM and WM×Belief in FD imply that demand consistency for WW and WM pairs in FD increases with higher beliefs, while it remains constant for MM pairs.[18] We summarize these new insights in the following result.

Result 4. *(A) In AN, belief accuracy and demand consistency of men and women do not differ. (B) In FD, irrespective of the gender of their counterpart, demands of women are more inconsistent than those of men. Women's demands become more consistent the more they expect the other to demand.*

The result shows that the less women expect the other to demand in FD, the less they demand themselves, suggesting that they are reluctant to exploit a low expected demand of a (known) opponent. Importantly, their higher demand inconsistency in FD apparently enables women to avoid disagreement when bargaining with other women. Men and women do not differ in their ability to estimate the other's behavior, but women display more cautious behavior by responding less to their own beliefs, thereby leaving a larger share on the bargaining table that helps to avoid disagreement. Given the observed general difficulty to estimate other's demand

---

[18]For instance, with a belief of 6000, women in a same gender pair would demand 8800 (13000 - 0.7 * 6000), which would result in 1200 left on the table. With a belief of 8000, they would demand 7400 (13000 - 0.7 * 8000), which would result in 600 left on the table. A belief of 10000 would give a demand of 6000, which would leave nothing on the table. With a belief larger than 10000, the regression coefficients predict that more would be demanded than what is left by the belief. From Figure 3, however, it is clear that this case does not occur. Women's beliefs in FD are at 10000 or lower.

accurately, women's lower consistency in demands gives more margin for agreement, whereas men's mostly consistent demands appear to aggressive, often leading to disagreement.

In sum, the relation between beliefs and demands, and particularly demand inconsistency, seems to be at the core of the observed gender differences in bargaining outcomes. Importantly, the results on demand consistency are robust to the inclusion of elicited risk attitudes of ego and alter (and other ego and alter controls) in the regressions, suggesting that risk aversion is not the relevant factor behind gender differences in demand consistency. Specifically, for women the Spearman correlation between risk aversion and demand consistency is only 0.001 ($p = 0.988$; $N = 175$) indicating that for women risk aversion is not the driver of inconsistent demands. (for details, see Appendix A.2). Rather, demand consistency seems to be a valid concept that captures behavior beyond mere risk aversion.

5. Conclusion

In this study we report on a controlled bargaining experiment in the field and analyze how gender and gender pairing influence bargaining outcomes. We conducted an artefactual field experiment in rural Uganda, using a one-shot Nash demand game (NDG) where two participants bargained over the division of a resource. During the experiment, negotiations between the two bargainers took place by simultaneously and independently demanding a share of the available resource. Bargaining pairs could be composed of two men (men–men pairs, MM), two women (women–women pairs, WW), or one man and one woman (mixed-gender pairs, WM and MW). To analyze the effect of gender and gender pairing on bargaining behavior, we exogenously varied identity disclosure. Depending on the information condition, participants either had information about the identity of their counterpart (full disclosure, FD) or they did not (anonymous, AN). The experiment was complemented with an elaborate socioeconomic survey to elicit important characteristics of the participants and their households.

Our results show that, while gender does not influence individual earnings in AN, substantial gender differences exist in FD. Specifically, we saw that, in FD, women earn more when paired with a woman than when paired with a man. Men's final earnings, on the contrary, do not depend on the gender of their counterpart. However, when the bargaining counterpart is a woman, men earn less than women. This indicates that participants in female same-gender

pairs are able to secure higher final earnings than either side in mixed-gender pairs. We further saw that earnings in case of agreement ('agreed earnings') do not exhibit gender differences, and do not depend on the gender of the counterpart for either men or women.

The absence of gender differences in agreed earnings indicates that the observed gender differences in final earnings are likely related to differences in disagreement rates between gender pairs. Indeed, in FD, the probability of disagreement is more than 30 percentage points lower among WW pairs than among mixed-gender pairs, while disagreement is equally likely in MM pairs and mixed-gender pairs. The lower disagreement rate of WW pairs is in line with their higher final earnings, indicating that it is precisely this lower disagreement rate that enables women in same-gender pairs to secure higher earnings than those in mixed-gender pairs.

In line with the observed gender differences in disagreement rates, we saw that, in FD, men demand more than women, but only when they are paired with a woman. If men in a mixed-gender pair demand more, while women behave the same in a mixed-gender and in a same-gender pair, more disagreements are likely to occur in mixed-gender pairs. Hence, men's higher demand in a mixed-gender pair causes disagreement which leaves not only themselves, but also those women they are matched with, worse off.

Based on the observed differences in bargaining outcomes, we proceeded to consider the role of beliefs in inducing such differences. Beliefs themselves do not differ across gender pairings, which suggests that gender differences in demands, disagreement, and earnings could be explained by potential gender differences in the relation between beliefs and demands. We defined the concepts of belief accuracy and demand consistency to analyze this relation. A belief is accurate if it is equal to the opponent's demand, and a demand is consistent if one's own belief and demand add up to exactly the available resource. Our results show that, in FD, irrespective of the gender of their counterpart, women have more inconsistent demands than men have, while belief accuracy does not differ across gender pairings.

The relation between beliefs and demands, and particularly demand consistency, seems to be at the core of the observed gender differences in bargaining outcomes. Women's lower demand consistency apparently enables them to avoid disagreement when in a female same-gender pair. While men and women do not differ in their ability to estimate the other's behavior,

24

women display more cautious behavior by leaving a larger share on the bargaining table that can help avoid disagreement. As men and women tend to underestimate the opponent's demand, disagreement is likely if both bargaining sides follow their beliefs when making their demands. Therefore, women's lower demand consistency is beneficial to them because it leaves more margin for agreement.

Moreover, women's demands become more consistent the more they expect the other to demand, while men's demand consistency is constant. Thus, women appear to be less willing to exploit a low expected demand of the opponent than men. This lower willingness of women to exploit an opponent might be linked to fairness preferences, consistent with some existing evidence showing that women have stronger preferences for equality (Andreoni and Vesterlund 2001, Cox 2002, Croson and Gneezy 2009). On the other hand, women's reluctance to exploit an opponent is also in line with the notion that women are less competitive than men (Gneezy et al. 2003, Niederle and Vesterlund 2007, Croson and Gneezy 2009, Niederle and Vesterlund 2011).

A more detailed analysis of the drivers of demand inconsistency, focusing in particular on confidence in own beliefs, competitiveness, and fairness preferences provides an interesting avenue for further research. We saw that results on demand consistency are robust to the inclusion of ego and alter controls, indicating that risk aversion is not the relevant factor behind gender differences in demand consistency. Rather, demand consistency was shown to be a valid concept that captures behavior beyond mere risk aversion. It remains to be analyzed though whether the behavior that is captured by demand consistency relates more to (over)confidence, competitiveness, or fairness preferences.

References

Adenzato, M., Brambilla, M., Manenti, R., De Lucia, L., Trojano, L., Garofalo, S., Enrici, I., and Cotelli, M. (2017). Gender differences in cognitive theory of mind revealed by transcranial direct current stimulation on medial prefrontal cortex. *Scientific Reports*, 7(1):1–9.

Altmann, S., Dohmen, T., and Wibral, M. (2008). Do the reciprocal trust less? *Economics Letters*, 99(3):454–457.

Andreoni, J. and Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics*, 116(1):293–312.

Armantier, O. and Treich, N. (2013). Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review*, 62:17 – 40.

Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4):14721503.

Azrieli, Y., Chambers, C. P., and Healy, P. J. (2020). Incentives in experiments with objective lotteries. *Experimental Economics*, 23:1–29.

Babcock, L. and Laschever, S. (2003). *Women don't ask: Negotiation and the gender divide*. Princeton, NJ: Princeton University Press.

Baron-Cohen, S., O'riordan, M., Stone, V., Jones, R., and Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with asperger syndrome or high-functioning autism. *Journal of autism and developmental disorders*, 29(5):407–418.

Blanco, M., Engelmann, D., Koch, A. K., and Normann, H.-T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, 13(4):412–438.

Bowles, H. R., Babcock, L., and Lai, L. (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes*, 103(1):84–103.

Brackett, M. A. and Salovey, P. (2006). Measuring emotional intelligence with the mayer-salovery-caruso emotional intelligence test (msceit). *Psicothema*, 18:34–41.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics*, 90(3):414–427.

Cardenas, J.-C., Dreber, A., von Essen, E., and Ranehill, E. (2012). Gender differences in competitiveness and risk taking: Comparing children in colombia and sweden. *Journal of Economic Behavior and Organization*, 83(1):11–23.

Castillo, M., Petrie, R., Torero, M., and Vesterlund, L. (2013). Gender differences in bargaining outcomes: A field experiment on discrimination. *Journal of Public Economics*, 99:35–48.

Charness, G. and Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, 83(1):50–58.

Costa, P. T. and McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO personality inventory. *Psychological Assessment*, 4(1):5.

Cox, J. C. (2002). Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. In *Experimental Business Research*, pages 331–350.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2):448–474.

Croson, R. T. (2007). Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. *Economic Inquiry*, 45(2):199–216.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2008). Representative trust and reciprocity: Prevalence and determinants. *Economic Inquiry*, 46(1):84–90.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2012). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies*, 79(2):645–677.

Dohmen, T. J., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2005). Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. IZA Discussion Paper, No. 1730. Bonn: Institute of Labor Economics.

Eckel, C. and Grossman, P. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry*, 39(2):171–188.

Eckel, C. C. and Grossman, P. J. (2008). Chapter 113 men, women and risk aversion: Experimental evidence. In Plott, C. R. and Smith, V. L., editors, *Handbook of Experimental Economics Results*, volume 1, pages 1061 – 1073. Elsevier.

Eriksson, K. H. and Sandberg, A. (2012). Gender differences in initiation of negotiation: Does the gender of the negotiation counterpart matter? *Negotiation Journal*, 28(4):407–428.

Fischbacher, U. and Gachter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review*, 100(1):541–56.

Gächter, S. and Renner, E. (2010). The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, 13(3):364–377.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5):1637–1664.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.

Handschuch, C. and Wollni, M. (2016). Traditional food crop marketing in sub-Saharan Africa: Does gender matter? *The Journal of Development Studies*, 52(3):343–359.

Harsanyi, J. C. (1962). Bargaining in ignorance of the opponent's utility function. *Journal of Conflict Resolution*, 6(1):29–38.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). Most people are not weird. *Nature*, 466(7302):29–29.

Hernandez-Arenaz, I. and Iriberri, N. (2018). Women ask for less (only from men): Evidence from bargaining in the field. *Journal of Economic Behavior & Organization*, 152:192 – 214.

Hernandez-Arenaz, I. and Iriberri, N. (2019). *Oxford Research Encyclopedia of Economics and Finance*, chapter A Review of Gender Differences in Negotiation. Oxford University Press.

Houweling, T. A., Kunst, A. E., and Mackenbach, J. P. (2003). Measuring health inequality among children in developing countries: does the choice of the indicator of economic status matter? *International journal for equity in health*, 2(1):8.

Keaveny, T. J. and Inderrieden, E. J. (2000). Gender differences in pay satisfaction and pay expectations. *Journal of Managerial Issues*, pages 363–379.

Kolb, D. M. (2009). Too bad for the women or does it have to be? Gender and negotiation research over the past twenty-five years. *Negotiation Journal*, 25(4):515–531.

Kray, L. J. and Thompson, L. (2004). Gender stereotypes and negotiation performance: An examination of theory and research. *Research in organizational behavior*, 26:103–182.

Kugler, K. G., Reif, J. A., Kaschner, T., and Brodbeck, F. (2018). Gender differences in the initiation of negotiations: A meta-analysis. *Psychological bulletin*, 144(2):198222.

Mazei, J., Hffmeier, J., Freund, P. A., Stuhlmacher, A. F., Bilke, L., and Hertel, G. (2015). A meta-analysis on gender differences in negotiation outcomes and their moderators. *Psychological Bulletin*, 141(1):85104.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological bulletin*, 126(3):424.

Nash, J. F. (1953). Two-person cooperative games. *Econometrica*, 21(1):128–140.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Niederle, M. and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3(1):601–630.

Rutström, E. E. and Wilcox, N. T. (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior*, 67(2):616–632.

Shan, W., Keller, J., and Joseph, D. (2019). Are men better negotiators everywhere? A meta-analysis of how gender differences in negotiation performance vary across cultures. *Journal of Organizational Behavior*, 40(6):651–675.

Small, D., Gelfand, M., Babcock, L., and Gettman, H. (2007). Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation. *Journal of personality and social psychology*, 93(4):600–613.

Solnick, S. (2001). Gender differences in the ultimatum game. *Economic Inquiry*, 39(2):189–200.

Vyas, S. and Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health policy and planning*, 21(6):459–468.

Walther, O. and Trémolières, M. (2019). Women and trade networks in West Africa.

World Bank (2012). World development report 2012: Gender equality and development. Washington, DC: World Bank.

Appendix

A. Socioeconomic Characteristics

Table A.1: Socioeconomic Characteristics

	Women			Men			
	N	mean	st.dev.	N	mean	st.dev.	p-value[+]
Risk	115	2.078	0.84	110	1.909	0.87	0.080
Age	115	42.99	14.4	110	39.68	13.7	0.081
Education	115	5.226	3.57	109	6.945	3.37	0.000
Wealth	115	-0.332	1.91	110	0.061	2.08	0.088
Agree	115	-0.009	1.56	110	0.042	1.82	0.661
Trust	115	-0.245	1.29	110	0.176	1.29	0.003
Friend	115	0.629	0.43	109	0.677	0.40	0.455

[+] two-sided p-value of a Mann-Whitney U test. Risk is a self-reported measure on a scale from 1 (low risk aversion) to 4 (high risk aversion). Education measures years of schooling. Wealth, Agree, and Trust are first components of a principal component analysis. Friend denotes the fraction of people in the community that a respondent calls a friend.

A.1. Composition of Wealth, Trust, and Agreeableness Index

A.1.1. Wealth Index

First, we performed a principal component analysis with the full set of asset indicators data elicited in the survey, restricted to four components:[19]

pca owned rooms floorearth floorearthdung floorcement waterprivate waterpublictap waterhole waterprowell waterunprowell waterriver watergravity covpitprivate covpitshared uncovpit electricity lantern cattleindigenous cattleexocross goats sheep pigs land bicycles motorvehicles generator stove sofas beds radios televisions jewellerywatches phones computers HHappliances storage stalls watercans irritubes insecpumps pulpchines grinders handthreshers wheelbarrows animalplough, cor components(4)

---

[19]An explanation of variable names and related survey questions can be found in the on-line Supplementary Information.

Although the eigenvalue of this PCA is very high, the first component only explains 12% of the variation in the asset indicators data, which is on the low side of the range of variance accounted for by the first principal component in existing studies according to the literature reviewed (Houweling et al. 2003, Vyas and Kumaranayake 2006). The main reason for this might be that the correlation among the variables included is not very high and, thus, there is a lot of variation in the data. We rerun the PCA while excluding variables with eigenvectors lower than 10%:

pca rooms floorearthdung floorcement electricity lantern cattleindigenous cattleexocross goats land bicycles motorvehicles generator stove sofas beds radios televisions jewellerywatches phones HHappliances storage stalls watercans insecpumps pulpchines wheelbarrows animalplough, cor components(4)

As expected the variance explained (given by the eigenvalue) increases from 12% to 19%. The first component of this latter PCA is included as Wealth in Table A.1.

A.1.2. Trust Index

We performed a principal component analysis based on the three trust questions (as explained in section 2.3) in the socioeconomic survey:

pca q8a1rev q8a2 q8a3rev, cor components(3)

The first component of this PCA is included as Trust in Table A.1.

A.1.3. Agreeableness Index

We performed a principal component analysis based on the agreeableness questions (as explained in section 2.3) in the socioeconomic survey:

pca q8b1rev q8b2 q8b3rev q8b4 q8b5rev q8b6 q8b7 q8b8 q8b9 q8b10rev q8b11 q8b12rev q8b13 q8b14rev q8b15rev q8b16rev, cor components(4)

The first component of this PCA is included as Agree in Table A.1.

A.2. Risk Aversion

Risk aversion was elicited as a self-reported measure on a four-point scale, with a higher number indicating more risk aversion. The variable Risk in Table A.1 reports descriptive statistics and Figure A.1 displays the distribution of risk aversion by gender, with women's results in the left panel and men's results in the right panel. We see that for both women and men the mode is at 2. However, for men, the bar at 1 seems slightly higher than for women, while for women the bar at 3 seems slightly higher than for men. This difference indicates a slightly higher risk aversion of women in our sample (two-sided p-value of a Mann-Whitney U test $= 0.080$; $N = 225$), in line with prior studies.[20]
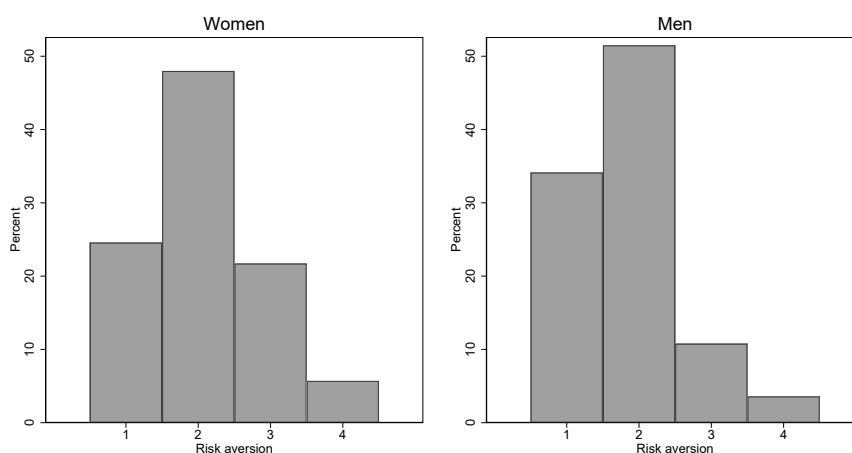
Figure A.1: Risk Aversion

The Spearman correlation between risk aversion and demand inconsistency is 0.107 (p-value $= 0.048$; $N = 342$). This indicates that higher risk aversion goes along with higher inconsistency of demands. However, a closer look shows that, while this Spearman correlation is 0.193 for men (p-value $= 0.013$; $N = 167$), it is only 0.001 for women (p-value $= 0.988$; $N = 175$). Further, for men, the Spearman correlation is 0.312 in AN (p-value $= 0.004$; $N = 83$), but only 0.035 in FD (p-value $= 0.753$; $N = 84$). This shows that the above conclusion, that higher risk aversion goes along with more inconsistent demands, strictly speaking only holds for men in AN. For men in FD, and for women in either AN or FD, risk aversion and inconsistency seem largely unrelated. In sum, demand consistency seems to be related to participants' risk aversion, but is far from being determined by it.

---

[20]All participants who understood the game are included here ($N = 225$), independently of any specific counterpart that they were matched with during the experiment.

B. Non-parametric Tests

Table B.1: Individual Earnings

	Women			Men			
	N	mean	st.dev.	N	mean	st.dev.	p-value[+]
a) Final Earnings							
AN	97	5041.24	3003.18	83	4746.99	2895.93	0.548
FD							
Woman	46	5543.48	2841.80	32	3937.50	2839.10	0.022
Man	32	3812.50	2705.28	52	4788.46	2939.50	0.145
p-value[++]		0.011			0.220		0.229
b) Agreed Earnings							
AN	53	7566.04	1525.54	43	7302.33	1581.66	0.627
FD							
Woman	30	7433.33	1381.74	11	7636.36	1433.37	0.588
Man	11	7272.73	1618.08	26	7576.92	1205.76	0.656
p-value[++]		0.885			0.730		0.693

Columns indicate ego's gender. In panel b) rows show alter's gender. [+] indicates the two-sided p-value of a Mann-Whitney U test by gender of ego. [++] indicates the two-sided p-value of a Mann-Whitney U test by gender of alter. Number in the bottom right corner indicates two-sided p-value of a Mann-Whitney U test between WW and MM.

Table B.2: Disagreement Rates

	AN		FD		p-value
	N	%	N	%	
WW	26	42.31	23	34.78	0.590
MM	19	47.37	26	50.00	0.862
Mix	45	48.89	32	65.63	0.145
p-value[+]	0.864		0.076		

[+] Based on Chi-square test

Table B.3: Demands

	Women			Men			
	N	mean	st.dev.	N	mean	st.dev.	p-value[+]
AN	97	8577.32	2249.24	83	8650.60	2461.58	0.791
FD							
Woman	46	8282.61	2051.16	32	8968.75	1596.05	0.044
Man	32	8562.50	1644.88	52	8913.46	2261.62	0.995
p-value[++]		0.223			0.355		0.159

Columns indicate ego's gender. In panel b) rows show alter's gender. [+] indicates the two-sided p-value of a Mann-Whitney U test by gender of ego. [++] indicates the two-sided p-value of a Mann-Whitney U test by gender of alter. Number in the bottom right corner indicates two-sided p-value of a Mann-Whitney U test between WW and MM.

Table B.4: Beliefs

	Women			Men			
	N	mean	st.dev.	N	mean	st.dev.	p-value[+]
AN	97	6216.49	2227.79	83	6493.98	2297.26	0.162
FD							
Woman	46	6826.09	2079.58	32	6531.25	1523.67	0.394
Man	32	6687.50	1803.89	52	6817.31	2176.14	0.492
p-value[++]		0.706			0.184		0.801

Columns indicate ego's gender. In panel b) rows show alter's gender. [+] indicates the two-sided p-value of a Mann-Whitney U test by gender of ego. [++] indicates the two-sided p-value of a Mann-Whitney U test by gender of alter. Number in the bottom right corner indicates two-sided p-value of a Mann-Whitney U test between WW and MM.

Table B.5: Belief Accuracy and Demand Consistency

	Woman			Man			
	N	mean	st.dev.	N	mean	st.dev.	p-value[+]
a) Belief Accuracy							
AN	97	-2381.44	3120.64	83	-2132.53	2937.44	0.768
FD							
Woman	46	-1456.52	2979.24	32	-2031.25	2348.43	0.269
Man	32	-2281.25	2666.81	52	-2096.15	3635.39	0.251
p-value[++]		0.166			0.390		0.509
b) Demand Consistency							
AN	97	1206.19	2277.33	83	855.42	2078.53	0.189
FD							
Woman	46	891.30	1608.70	32	500.00	1047.27	0.250
Man	32	750.00	1414.21	52	269.23	769.91	0.077
p-value[++]		0.623			0.291		0.012

Columns indicate ego's gender. In panel b rows show alter's gender. [+] indicates the two-sided p-value of a Mann-Whitney U test by gender of ego. [++] indicates the two-sided p-value of a Mann-Whitney U test by gender of alter. Number in the bottom right corner indicates two-sided p-value of a Mann-Whitney U test between WW and MM.

C. Additional Regression Tables

C.1. Earnings

Table C.1: The Effect of Gender on Individual Earnings in AN

	Final Earnings		Agreed Earnings	
	(1)	(2)	(3)	(4)
Men	-424.6	-522.4	-296.7	-501.7
	(480.6)	(519.9)	(369.9)	(344.2)
Educ ego		-32.89		66.59
		(79.55)		(62.36)
Risk ego		-281.8		-755.7***
		(209.9)		(233.5)
Trust ego		28.00		-217.8*
		(199.4)		(132.0)
Wealth ego		138.5		-37.44
		(123.8)		(80.21)
Age ego		0.369		15.10
		(15.84)		(10.11)
Constant	5889.4***	6701.2***	7735.1***	8233.9***
	(397.2)	(1250.2)	(231.5)	(822.8)
Observations	180	179	96	95
R^2	0.063	0.079	0.014	0.181
Controls	No	Yes	No	Yes

Linear regressions. Dependent variables are final earnings (models 1 & 2) and earnings in the case of agreement (models 3 & 4). Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Table C.2: The Effect of Gender Pairings on Individual Earnings in FD

	Final Earnings		Agreed Earnings	
	(1)	(2)	(3)	(4)
MM	-751.0	-244.2	149.7	31.18
	(764.4)	(1135.7)	(140.2)	(297.4)
WM	-1728.4***	-1788.8***	-161.0	-370.4
	(647.0)	(655.5)	(551.3)	(553.6)
MW	-1603.4***	-1499.7**	202.6	472.2
	(614.9)	(589.0)	(333.8)	(368.2)
Educ ego		-64.02		-32.05
		(97.07)		(84.05)
Risk ego		41.64		21.47
		(351.9)		(286.2)
Trust ego		-118.4		23.83
		(208.8)		(205.0)
Wealth ego		112.1		-7.939
		(117.8)		(123.8)
Age ego		28.91*		12.22
		(15.78)		(19.60)
Educ alter		-3.589		99.45
		(104.0)		(93.18)
Risk alter		97.34		115.6
		(351.9)		(257.2)
Trust alter		-96.35		124.5
		(247.2)		(169.5)
Wealth alter		143.4		45.39
		(92.62)		(113.5)
Age alter		22.42		5.865
		(21.08)		(18.79)
Constant	5527.9***	3633.8*	7414.5***	6026.8***
	(656.5)	(2143.5)	(167.6)	(1024.4)
Observations	162	162	78	78
R^2	0.056	0.107	0.008	0.110
Controls	No	Yes	No	Yes
Wald tests (p-values)				
MM - WM	0.340	0.200	0.563	0.409
MM - MW	0.396	0.299	0.887	0.359

Linear regressions. Dependent variables are final earnings (models 1 & 2) and earnings in the case of agreement (models 3 & 4). Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

C.2. Disagreement

Table C.3: The Effect of Gender Pairings on Pairwise Disagreement in AN

	Disagreement	
	(1)	(2)
WW	-0.117	-0.168
	(0.119)	(0.128)
MM	-0.0268	-0.140
	(0.152)	(0.181)
Educ alter		0.0393*
		(0.0204)
Risk alter		0.0696
		(0.0884)
Trust alter		-0.000466
		(0.0525)
Wealth alter		-0.0394
		(0.0315)
Age alter		0.00146
		(0.00465)
Educ ego		-0.00381
		(0.0274)
Risk ego		-0.0550
		(0.0661)
Trust ego		-0.0638
		(0.0507)
Wealth ego		-0.0113
		(0.0371)
Age ego		0.00416
		(0.00490)
Constant	0.372***	-0.0895
	(0.0881)	(0.613)
Observations	90	89
R^2	0.069	0.165
Controls	No	Yes
Wald tests (p-values)		
WW - MM	0.608	0.895

Linear regressions. Dependent variable is disagreement. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Table C.4: The Effect of Gender Pairings on Pairwise Disagreement in FD

	Disagreement	
	(1)	(2)
WW	-0.308***	-0.347***
	(0.0998)	(0.116)
MM	-0.156	-0.263
	(0.182)	(0.236)
Educ alter		0.0201
		(0.0267)
Risk alter		-0.0403
		(0.0872)
Trust alter		-0.00912
		(0.0524)
Wealth alter		0.0205
		(0.0322)
Age alter		-0.00438
		(0.00363)
Educ ego		-0.00441
		(0.0269)
Risk ego		-0.00291
		(0.108)
Trust ego		0.0860
		(0.0619)
Wealth ego		-0.0389
		(0.0275)
Age ego		-0.00582
		(0.00726)
Constant	0.658***	1.093***
	(0.112)	(0.406)
Observations	81	81
R^2	0.064	0.163
Controls	No	Yes
Wald tests (p-values)		
WW - MM	0.278	0.702

Linear regressions. Dependent variable is disagreement. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Table C.5: The Effect of Gender Pairings on Pairwise Disagreement in AN (Probit)

	Disagreement	
	(1)	(2)
disagreement		
WW	-0.315	-0.480
	(0.342)	(247.0)
MM	-0.0716	-0.412
	(0.441)	(233.1)
Educ alter		0.112
		(106.0)
Risk alter		0.191
		(135.7)
Trust alter		-0.00686
		(107.5)
Wealth alter		-0.117
		(134.8)
Age alter		0.00447
		(3.124)
Educ ego		-0.0194
		(51.59)
Risk ego		-0.179
		(415.7)
Trust ego		-0.185
		(215.2)
Wealth ego		-0.0289
		(190.0)
Age ego		0.0112
		(6.381)
Constant	-0.329	-1.542
	(0.249)	(137.0)
Observations	90	89
R^2		
Controls	No	Yes
Wald tests (p-values)		
WW - MM	0.634	0.996

Probit regressions. Dependent variable is disagreement. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: [*] $p < 0.10$, [**] $p < 0.05$, [***] $p < 0.01$ indicate two-sided significance levels.

Table C.6: The Effect of Gender Pairings on Pairwise Disagreement in FD (Probit)

	Disagreement	
	(1)	(2)
disagreement		
WW	-0.792***	-0.970
	(0.280)	(12.28)
MM	-0.403	-0.765
	(0.517)	(58.06)
Educ alter		0.0618
		(3.807)
Risk alter		-0.138
		(10.15)
Trust alter		-0.0373
		(6.302)
Wealth alter		0.0625
		(6.650)
Age alter		-0.0117
		(0.870)
Educ ego		-0.0155
		(3.261)
Risk ego		-0.0225
		(6.354)
Trust ego		0.245
		(4.976)
Wealth ego		-0.123
		(7.122)
Age ego		-0.0173
		(0.969)
Constant	0.410	1.747
	(0.325)	(98.91)
Observations	81	81
R^2		
Controls	No	Yes
Wald tests (p-values)		
WW - MM	0.326	0.998

Probit regressions. Dependent variable is disagreement. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

C.3. Demands

Table C.7: The Effect of Gender on Demands in AN

	(1)	(2)
Men	127.1	-74.73
	(320.1)	(322.4)
Educ ego		80.75
		(54.35)
Risk ego		-356.9
		(241.4)
Trust ego		20.98
		(167.5)
Wealth ego		-133.8*
		(74.35)
Age ego		-17.82
		(14.13)
Constant	8223.0***	9187.3***
	(345.8)	(936.3)
Observations	180	179
R^2	0.021	0.079
Controls	No	Yes

Linear regressions. Dependent variable is ego's demand. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Table C.8: The Effect of Gender Pairings on Demands in FD

	(1)	(2)
MM	707.2	697.4
	(438.7)	(620.0)
WM	330.1	335.0
	(350.9)	(321.2)
MW	736.4***	844.5**
	(269.2)	(394.8)
Educ ego		-4.385
		(60.66)
Risk		-37.33
		(244.4)
Trust ego		-22.66
		(195.4)
Wealth ego		-117.6**
		(47.40)
Age ego		4.619
		(17.45)
Educ alter		51.06
		(64.73)
Risk alter		174.8
		(187.4)
Trust alter		92.26
		(139.6)
Wealth alter		-28.16
		(70.96)
Age alter		1.517
		(11.45)
Constant	7983.1***	7060.2***
	(364.6)	(1486.7)
Observations	162	162
R^2	0.028	0.052
Controls	No	Yes
Wald tests (p-values)		
MM - WM	0.537	0.603
MM - MW	0.958	0.834

Linear regressions. Dependent variable is ego's demand. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

C.4. The Role of Beliefs

Table C.9: The Effect of Gender on Beliefs in AN

	(1)	(2)
Men	173.0	80.49
	(275.7)	(301.1)
Educ ego		-3.303
		(51.24)
Risk ego		-49.58
		(180.9)
Trust ego		61.82
		(163.6)
Wealth ego		260.4***
		(72.13)
Age ego		19.49
		(14.06)
Constant	6887.3***	6259.8***
	(328.8)	(1071.9)
N	180	179
R^2	0.0535	0.128
Controls	No	Yes

Linear regressions. Dependent variable is ego's belief. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Table C.10: The Effect of Gender Pairings on Beliefs in FD

	(1)	(2)
MM	-149.8	-302.5
	(561.7)	(655.2)
WM	-231.3	-197.4
	(543.2)	(557.5)
MW	-387.6	-512.8
	(436.1)	(527.2)
Educ ego		47.76
		(66.10)
Risk ego		133.5
		(236.3)
Trust ego		57.46
		(213.0)
Wealth ego		166.8*
		(88.10)
Age ego		-7.397
		(15.66)
Educ alter		-30.17
		(71.26)
Risk alter		-169.7
		(207.3)
Trust alter		-195.8
		(164.3)
Wealth alter		45.27
		(60.40)
Age alter		-2.609
		(13.05)
Constant	7378.8***	7911.6***
	(534.8)	(2125.8)
N	162	162
R^2	0.0274	0.0925
Controls	No	Yes
Wald tests (p-values)		
MM - WM	0.879	0.866
MM - MW	0.623	0.748

Linear regressions. Dependent variable is ego's belief. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Table C.11: The Effect of Gender on Belief Accuracy and Demand Consistency in AN

| | Belief Accuracy | | Demand Consistency | |
	(1)	(2)	(3)	(4)
Women	-713.1	-586.6	-1378.5	-1357.7
	(694.5)	(776.8)	(1583.2)	(1418.8)
Belief	0.123	0.122	-0.664***	-0.692***
	(0.0864)	(0.0963)	(0.173)	(0.169)
Women × Belief	0.104	0.0937	0.182	0.220
	(0.125)	(0.138)	(0.210)	(0.189)
Educ ego		42.36		80.66*
		(41.94)		(47.75)
Risk ego		20.61		-403.8*
		(235.2)		(211.1)
Trust ego		168.1		58.28
		(119.8)		(105.4)
Wealth ego		-51.30		12.56
		(93.23)		(63.59)
Age ego		17.48*		-7.490
		(10.17)		(11.08)
Constant	7440.4***	6408.9***	13001.1***	13592.8***
	(626.8)	(1185.1)	(1331.8)	(1570.0)
Observations	180	179	180	179
R^2	0.051	0.070	0.312	0.359
Controls	No	Yes	No	Yes

Linear regressions. In Model 1 & 2, dependent variable is alter's demand. In Model 3 & 4, dependent variable is ego's demand. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Table C.12: The Effect of Gender Pairings on Belief Accuracy and Demand Consistency in FD

| | Belief Accuracy | | Demand Consistency | |
	(1)	(2)	(3)	(4)
WW	-2746.2	-2490.4	-2675.9**	-2578.6**
	(1981.7)	(2123.5)	(1185.8)	(1136.1)
WM	-943.7	-434.4	-3101.6*	-3270.6**
	(1890.9)	(1964.2)	(1640.3)	(1494.8)
MW	-1952.3	-1274.1	-1575.4	-1292.0
	(2150.2)	(2372.6)	(1204.8)	(1152.3)
Belief	-0.342	-0.328	-1.004***	-1.036***
	(0.234)	(0.256)	(0.113)	(0.0904)
WW $\times$ Belief	0.306	0.274	0.310**	0.321**
	(0.249)	(0.270)	(0.151)	(0.141)
WM $\times$ Belief	0.141	0.0881	0.395*	0.450**
	(0.235)	(0.250)	(0.223)	(0.207)
MW $\times$ Belief	0.229	0.127	0.209	0.188
	(0.282)	(0.305)	(0.158)	(0.147)
Educ ego		57.07		23.77
		(59.52)		(36.56)
Risk ego		174.8		30.07
		(190.7)		(150.0)
Trust ego		114.6		8.618
		(135.6)		(95.09)
Wealth ego		9.517		49.38
		(73.13)		(59.86)
Age ego		0.00787		-5.077
		(12.43)		(9.294)
Educ alter		-11.83		35.83
		(57.17)		(26.00)
Risk alter		-73.74		43.27
		(244.8)		(162.4)
Trust alter		-70.75		-63.87
		(195.5)		(72.81)
Wealth alter		-96.60*		18.95
		(54.35)		(44.74)
Age alter		4.108		0.0938
		(16.57)		(5.423)
Constant	11158.2***	10288.9***	15937.2***	15740.4***
	(1987.2)	(2143.8)	(893.9)	(1087.3)
Observations	162	162	162	162
R^2	0.079	0.103	0.676	0.688
Controls	No	Yes	No	Yes
Wald tests (p-values)				
WW - WM	0.144	0.185	0.840	0.718
WW - MW	0.512	0.325	0.422	0.317
WW$\times$Belief - WM$\times$Belief	0.360	0.386	0.766	0.619
WW$\times$Belief - MW$\times$Belief	0.663	0.419	0.575	0.421

Linear regressions. In Model 1 & 2, dependent variable is alter's demand. In Model 3 & 4, dependent variable is ego's demand. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at session level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

D.   Experimental Instructions

You will be paired with one other person in this room. Both of you will be asked to make a decision. Your decision as well as the decision of the other person will determine how much you can earn. These earnings depend on your own decision and the decision of the other person. Your earnings will be determined in the following way. In each pair we have two persons: person 1 and person 2. There are 16000 UGX on the table [put 16 notes of 1000 UGX on the table] and person 1 and person 2 can demand as much as they want of it. If the total person 1 and person 2 demand is not higher than the money on the table (that is 16000 UGX) each will get the amount demanded.

However, if the total is more than 16000 UGX none will get the amount demanded and person 1 will get 2000 UGX and person 2 will receive 2000 UGX. For example, imagine person 1 demanded 10,000 UGX and person 2 demanded 6,000 UGX. What do they demand in total? (16,000 UGX). Do we have enough on the table? (yes). As there is enough on the table each person will get what he/she demanded. Person 1 gets 10,000 UGX and person 2 gets 6,000 UGX.

Imagine now that person 1 demanded 11,000 UGX and person 2 demanded 7,000 UGX.

- What do they demand in total? (18,000 UGX).
- Do we have enough on the table? (no).

As there is NOT enough on the table person 1 would get 2,000 UGX and person 2 would get 2,000 UGX.

Let me check whether you understood [Ask the following questions in public and ask the participants to respond.]

- How much income would person 1 get if he demanded 5,000 UGX and person 2 demanded 11,000 UGX? (5,000 UGX). How much would person 2 get? (11,000 UGX)
- How much income would person 2 get if he demanded 8,000 UGX and person 1 demanded 11,000 UGX? (2,000 UGX). How much would person 1 get? (2,000 UGX)

It is important to remember that at the time you make your decision you do not know the decision of the person you are paired with. Similarly, the other person does not know your decision, when making

his/her own decision. You can of course have beliefs about what the other will demand. [Ask the following questions in public and ask the participants to respond.]

1. Imagine that person 1 believes that person 2 will demand 9,000 UGX. How much would person 1 get if he demanded 9,000 UGX as well? (2,000 UGX). How much would person 2 get? (2,000 UGX)

2. Imagine that person 1 believes that person 2 will demand 9,000 UGX. How much would person 1 get if he demanded 6,000 UGX? (6,000 UGX). How much would person 2 get? (9,000 UGX)

3. Imagine that person 2 believes that person 1 will demand 6,000 UGX. How much would person 2 get if he demanded 6,000 UGX as well? (6,000 UGX) . How much would person 1 get? (6,000 UGX)

4. Imagine that person 2 believes that person 1 will demand 6,000 UGX. How much would person 2 get if he demanded 11,000 UGX? (2,000 UGX) . How much would person 1 get? (2,000 UGX)

[Stick poster of decision card to the wall and distribute empty decision card] To make decisions, we will proceed in the following way. First, we will ask you to specify on the decision card what you believe the other would choose.[Use the poster to explain how to use the decision card] After this, you will be asked to mark your decision on your decision card. [Use the poster to explain how to use the decision card]

Pairing

You will make several decisions in which you will be paired with different persons in this room. At the end of today's programme we select one pair for your payment and you will get to know the identity of the other person in the selected pair and the other person in this pair will get to know your identity. However, at the moment when you will be asked to make a decision, you won't always know the identity of the person you are paired with.

In some pairs you wont know the identity of the other person, and neither will the other person know your identity. In this case the two boxes under YOU and Other person will be empty. The other person could be from the same village where you live or from another village. This will be indicated on the decision card [Show on the poster of the investment decision card where it will be indicated whether same/different village]. Semi-disclosure treatment: In other pairs, one person will know the identity of the person s/he is paired with, while the other person will not know the identity of the person s/he is

paired with. The person who will know the identity of the other person will find the name and photograph of the other person on his/her decision card. If you get to see your photograph on the decision card the other will know your identity and name. If your photograph/name does not appear on your decision card, the other will not know your identity. [Show on the poster of the decision card where they can find the names and photographs of both persons]

In other words, if you get to see a photograph and name in the box under Other person, you get to know the identity of the person you are paired with. If you see your photograph on your decision card, the other will know your identity and name. If your photograph/name does not appear on your decision card, the other wont know your identity.

For each of the pairs you are involved in you will receive a new decision card. You may make the same decision or you may make a different decision.

Control questions

We will now ask some questions to see whether you understood the instructions.

1. How much would you get if you demanded 10,000 UGX and the other person demanded 10,000 UGX as well? How much would the other person get?

2. How much would you get if you demanded 4,000 UGX and the other person demanded 12,000 UGX? How much would the other person get?

3. How much would you get if you demanded 8,000 UGX and the other person demanded 10,000 UGX? How much would the other person get?

4. How much would you get if you demanded 10,000 UGX and the other person demanded 6,000 UGX? How much would the other person get?

[For each of the questions, record on the control question card whether they answered it correctly. If the participant gave a wrong answer for at least one of the questions, ask him/her to have a careful look at it once more and ask what was not clear. Answer their questions as clearly and accurately as possible. If necessary, clarify the instructions; but not more than once. Write down additional comments if you think the participants did not get enough understanding. Retain their decision cards.]

Decisions

[Give each participant a pen.] If you have no further questions, we will now begin. Remember, there are no wrong decisions, so you should choose the option as you prefer. We emphasize that it is important that you make your decision in private. Do not show your decision card to the other participants. If you need assistance, please raise your hand so that one of us can come to assist you. Once you have made your decision, please fold the decision card and raise your hand so that we can come by to collect your decision card.

[The participants remain seated. We give decision card with pair no 1 to the participants. Clarify publicly the treatment (same/different village, anonymous/non-anonymous). After the participants have made their decision, they fold their decision card. When collecting the decision cards we check whether their answer is readable and consistent. Add comments if the participant was struggling (e.g. if he/she was helped with filling in the decision card). After all cards have been returned, we give them the decision card for pair no 2. Explain that it is a new pair and clarify publicly important elements such as the name/photograph of the involved participants (if relevant) including whose identity is known to whom, and whether they belong to the same village. Follow the same procedure for the other pairs. Make sure that distribution cards are distributed in the correct order 1 4.]

[When all participants have made their 4 decisions, the experiment is complete. Control that all decision cards have been returned. Collect pay-off table cards and remove poster]