Journal of Economic Behavior and Organization 205 (2023) 237–269

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

The effect of gender and gender pairing on bargaining: Evidence from an artefactual field experiment☆

Ben D'Exelle[a], Christine Gutekunst[b], Arno Riedl[c,*]

[a] University of East Anglia, School of Economics and School of International Development, United Kingdom
[b] Maastricht University, Department of Microeconomics and Public Economics, The Netherlands
[c] CESifo, IZA, and Maastricht University, Department of Microeconomics and Public Economics, The Netherlands

ARTICLE INFO

Article history:
Received 30 November 2020
Revised 7 August 2022
Accepted 16 October 2022

JEL classification:
C9
J16
O12

Keywords:
Bargaining
Gender
Gender pairing
Experiment

ABSTRACT

Men and women negotiate differently, which might create gender inequality in earnings from bargaining. We study the role of gender and gender pairing in bilateral bargaining, using an artefactual field experiment in rural Uganda, 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 and agreements, but only when identities and, thus, genders are disclosed. Women in same-gender pairs obtain higher final earnings than men and women in mixed-gender pairs, which is due to the lower likelihood of disagreement among women-only pairs. We identify gender differences in demands and demand inconsistency–the money left on the table once demands are corrected for beliefs about the counterpart's demand–as mechanisms behind the observed gender differences in bargaining outcomes. In addition, we find that gender differences in demand inconsistency are related to gender differences in education and a measure of expected generosity.

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 and Hernandez-Arenaz and Iriberri, 2019). Most evidence supports the view that women tend to be disadvantaged in bargaining, because they are more reluctant to initiate a negotiation or because they are less successful in negotiations (for the former, 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; for the latter, see, e.g., Kray and Thompson, 2004; Keaveny et al., 2000, and the meta-analysis of Mazei et al., 2015, of the psychological literature).

☆ We thank Jean-Marie Baland, Nathan Nunn and participants at the UNU-WIDER workshops in Helsinki and Namur for helpful comments, as well as participants at RES 2017 and EEA 2017, and seminar participants at UEA. We also thank Joshua Balungira, Alex Feuchtwanger and the Field Lab (https://thefieldlabuganda.com) for support with the field work. Ethical clearance for the study was obtained from the Uganda National Council for Science and Technology (UNCST): research registration number SS4089.

* Corresponding author.
E-mail addresses: b.dexelle@uea.ac.uk (B. D'Exelle), christine.a.gutekunst@gmail.com (C. Gutekunst), a.riedl@maastrichtuniversity.nl (A. Riedl).

https://doi.org/10.1016/j.jebo.2022.10.024
0167-2681/

Within this large literature, we identify two areas where more evidence is needed. First, several studies have documented that gender pairing matters for bargaining, but the evidence on how this matters is mixed. Second, men appear to be less likely to outperform women in negotiations in societies characterised by lower individualism. This is a conundrum as these societies are also characterized by larger gender inequality. In this paper, we aim to fill these gaps in the literature and link them by generating new experimental evidence.

Evidence on the role of gender pairing in bargaining is mixed. In an early study, using an ultimatum game, Eckel and Grossman (2001) observe that men are more likely to accept an offer from a woman than from a man, and that women-only pairs are more likely to agree on the division. 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. More recently, 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. 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. Finally, Hernandez-Arenaz and Iriberri (2018) observe that women demand less in a bargaining game in the field, but only when matched with a man.

Gender differences in bargaining vary considerably across cultures. Using a meta-analysis of 185 studies from 30 societies, Shan et al. (2019) documents that women do better than men in negotiations in cultures characterised by lower individualism. This is puzzling as these societies tend to be characterized by larger gender inequality and lower economic development. The cross-cultural study of gender differences in preferences of Falk and Hermle (2018) suggests that women's better performance might be due to smaller gender differences in preferences. Using representative samples in 76 countries, they find that gender differences in preferences correlate positively with economic development and negatively with gender inequality. If gender differences in preferences were responsible for women's worse bargaining performance compared to men, one would expect women to be in a better position where preference differences are lower.[1]

This reveals an important conundrum: How can women's better bargaining performance be reconciled with larger gender inequality? A possible reason for the puzzling evidence is that most studies focus on gender differences without taking account of gender pairing. Women might be particularly good at negotiating with other women, but be relatively unsuccessful in mixed-gender pairs. In our study we aim to contribute to answering this question by providing new experimental evidence on the effect of gender and gender pairing in bilateral bargaining (demands, agreements, earnings) in a society with large gender inequality. In addition, we also look at beliefs and socio-economic characteristics as potential mechanisms behind differences in bargaining behavior. To achieve this, we study bargaining behavior in a sample of men and women from rural Uganda. This is a society characterized by large gender inequality, low economic development and low individualism, i.e., a society where, based on the above evidence, we would expect women to out-perform men in negotiations.

Next to shedding light on this conundrum, our study also contributes to the literature by investigating potential mechanisms so far ignored. These mechanisms link demands and beliefs. Specifically, we explore whether belief inaccuracy—the difference between expected demand of the other and own actual demand—and demand inconsistency—the difference between the sum of belief and demand and the available resource—do explain gender differences in bargaining. We will see that demand inconsistency indeed plays a role. In addition, we also add new insights regarding the effect of the individual characteristics education and expected generosity in gender differences in bargaining.

We conduct an artefactual field experiment, where participants are randomly paired and bargain over a given resource in a Nash demand game (NDG) (Nash, 1953). Specifically, both participants in a pair simultaneously demand a share of a given amount of resources. If the sum of demands does not exceed the amount of resources available they receive their respective demands, otherwise they receive a low disagreement payoff. This game 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). It is also reminiscent of bargaining around locally traded goods and new opportunities or resources that are distributed among villagers, such as food or cash distributed by development projects organised by governmental or non-governmental organizations.

To investigate the role of gender and gender pairing in bargaining we use two treatments that vary the information about the bargaining opponent. In the baseline (anonymous) treatment, participants are not told whom they are 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 are disclosed which allows us to test whether the gender of the opponent influences bargaining and whether this differs between men and women. To avoid experimenter demand effects we do not actively tell participants the gender of the counterpart in this treatment. Rather, the disclosure of the bargainers identity gives participants the information necessary to identify the gender of their opponent. Arguably,

[1] Gender differences in risk preferences, preferences for competition and social preferences have been used to explain why men outperform women in negotiations in western societies. There is evidence from studies in western societies that men tend to be less risk averse than women (Eckel and Grossman, 2008; Croson and Gneezy, 2009a; Charness and Gneezy, 2012), have a stronger preference for competition (Gneezy et al., 2003; 2009; Niederle and Vesterlund, 2007; Cardenas et al., 2012) and tend be less generous than women (see, e.g., Bolton and Katok, 1995; Eckel and Grossman, 1996,1998; Cox and Deck, 2006; Konow et al., 2008 and Croson and Gneezy, 2009b, for a survey). In societies where these gender differences in preferences are less pronounced, gender differences in bargaining might be smaller and even reverse sign. Accordingly, studies conducted in developing countries did not find statistically significant gender differences in generosity (e.g. Binzel and Fehr (2013) in Cairo; Jakiela (2011) in Kenya; Gowdy et al. (2003) in Nigeria; Ligon and Schechter (2012) in Paraguay; Ado and Kurosaki (2014) in Jakarta). D'Exelle and Riedl (2019) found in rural Nicaragua that women tend to share even less than men, even after controlling for their larger friendship networks.

besides gender, other characteristics are disclosed when lifting anonymity. Therefore, 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.

Our results can be summarized as follows. First, we find gender differences in final earnings and agreements, but only when the identities, and thus gender, of the paired participants are disclosed. Specifically, when paired with a woman, women obtain higher earnings than men and than when paired with a man. Second, these gender differences in earnings are the result of a lower likelihood of disagreement among women-only pairs compared to mixed-gender pairs. Third, we do not find any gender differences in beliefs. Fourth, we identify gender differences in demands and demand inconsistency, that is, the money left on the bargaining table once demands are corrected for the beliefs regarding the bargaining partner's demand, as possible mechanisms behind the observed gender differences in agreements. Moreover, we find that gender differences in the individual characteristics education and a measure of expected generosity affect the demand inconsistency mechanism.

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 and disagreement rates. Section 4 explores potential mechanisms, and Section 5 concludes.

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 receive their respective demand. Otherwise, each player gets ex ante defined and known disagreement earnings $d_i = d_j = d$.[2] 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 actual bargaining efficiency is not guaranteed because, given the high strategic uncertainty, it is not obvious that demands and beliefs are mutually consistent. Thus, disagreements or inefficient agreements can be expected.

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.[3] Treatment AN allows us to investigate the effect of own gender in bargaining when not knowing the gender

[2] Generally, d_i and d_j can differ, however, in our experiment both players have symmetric disagreement earnings.

[3] 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.

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. Participants inferred the gender of their opponents from the pictures. An alternative approach would be to actively tell participants the gender of the counterpart, thereby risking the introduction of an experimenter demand bias. The participants could then suspect that we are interested in gender, and potentially adjust their behavior. For example, men might demand less if matched with a woman, if they think that we care about gender equality. Our approach has the advantage that it avoids such potential experimenter demand effects. At the same time, however, a number of other characteristics are disclosed in addition to gender.[4] This might confound the estimated effect of gender, if these characteristics correlate with both gender and bargaining behaviour. To minimize this risk, we followed the approach commonly used to take passport photographs: no smiling, only face shown, no clothes visible, etc. In addition, as many of the participants know each other—they live in the same small village—they might also know less visible characteristics of their opponent, such as wealth, age, education, trust, risk aversion, and whether they are friends. Therefore, in the analysis we control for potential confounding factors with the help of detailed information from a socioeconomic survey.

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.[5]

Further, we asked a variety of questions to elicit trust, expected fairness, expected generosity, and risk aversion (Dohmen et al., 2008; 2012; Altmann et al., 2008). Regarding trust, we asked the respondents whether they would say that most people can be trusted, or that they cannot be too careful in dealing with people. For expected fairness, the question was whether the respondent thought that people would try to take advantage of them if they got a chance, or that they would be fair. For expected generosity, we asked if most of the time people tried to be helpful, or if they are mostly just looking out for themselves. All three questions had three answer options, equivalent to 'agreeing fully,' 'agreeing somewhat,' and 'disagreeing fully.'[6] 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).

2.4. Implementation of the survey and experiment

The study was conducted in the Sironko District and Lower Bulambuli District in eastern Uganda. This area has a population size of around 300,000, most of which depend on agriculture, on relatively small land holdings (average land holdings are about 1.5 acres). Most people are from the Bagisu ethnic group (Verschoor et al., 2016). Jackson (2013) presents a clear description of the relations between men and women in this area. Life of the latter is characterized by male control and high domestic violence. However, while women are controlled by their father and once married by their husband, they do have considerable autonomy to marry, divorce and remarry. The latter makes men anxious as their reputation and perceived masculinity depend on them being married. People living in rural villages in this area bargain over goods or resources on a daily basis, such as locally traded goods, and new opportunities or resources that are offered by development projects organised by governmental or non-governmental organizations.

Across the study area, we selected a sample of participants, using a multi-cluster sampling approach. In the first step, we selected a random set of 24 villages. Villages in this area are small-scale, close-knit communities of the type that is common in the rural areas of most Sub-Saharan African countries. Most villages have between 30 and 60 households. In each of the villages, we selected a random sample of 25–30 households and from each of these selected households, one

[4] Laboratory experiments have found that attractiveness may affect behavior in the ultimatum game (Solnick and Schweitzer, 1999), trust game (Wilson and Eckel, 2006) and public goods game (Andreoni and Petrie, 2008). A recent field experiment showed that women receive higher discounts at produce markets in Israel and that the discounts further increase with their perceived attractiveness (Ruffle et al., 2022). In our setting, this should only be a problem if it is correlated with gender. Unfortunately, we did not measure attractiveness, and hence cannot test whether it correlates with gender. However, we can and do control for a number of other (un)observable characteristics. Moreover, Solnick and Schweitzer (1999) did not find an interaction effect between gender and attractiveness and offers made in an ultimatum game.

[5] 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 them to remember all their social ties.

[6] 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.'

adult was chosen at random to participate in the survey and experiment. Most households had only two adults. Thanks to the support of and collaboration with village leaders, attrition was very small, as nearly all invited households attended the experimental sessions. Overall 199 participants took part in the experiment, distributed across 14 sessions with between 10 and 20 participants.[7,8] The household survey was conducted in private 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 with different opponents in the AN condition and two decisions with other different opponents in the FD condition.[9] This allows us to observe variation in behavior of the same 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, no feedback was provided 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 tasks described on the decision cards that asked participants for the amount that they expected the other to demand, and the amount that they demanded for themselves, respectively. The resource available was 16,000 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. No further validity check, such as a restriction on the sum of a participant's belief and demand, was imposed. Demands were incentivized by informing participants that, at the end of the experiment, one of their decisions would be selected at random to be paid out. This protocol of randomly selecting one single decision for payment was followed to avoid hedging between different decisions (see Azrieli et al., 2018; 2020, for theoretical arguments and empirical evidence in favor of this procedure). Belief elicitation was not incentivized to avoid hedging between the demand and belief elicitation tasks.[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 F). If participants were unable to answer the control questions correctly further clarifications were given in private. If they were still unable to answer the control questions correctly, 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.[11] 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 made in private. Payments 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.[12] 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.

3. Results: Earnings and agreement

In the following, we describe the role of gender and gender pairing in determining individual earnings and pairwise disagreement outcomes.

[7] The reported experiment was part of a larger study. In this study, overall, there were 28 villages and 28 experimental sessions conducted. In 14 of these sessions the game used in this paper was played. These 14 sessions drew on participants from 24 villages. As participants from one village were spread across two to three sessions, some of the survey participants from these 24 villages were assigned to other sessions that were not part of the game analyzed in this paper. This implies that the number of participants from one village that took part in the game analyzed here was usually lower than the 25–30 mentioned in the main text (see Appendix Tables A.2 and A.3 for details).

[8] Camerer et al. (2016) report on a replication study of laboratory experiments published in two top-tier economic journals and found indications that a sufficient sample size is important for the successful replication of results. We believe that our sample size of 199 participants together with the statistically powerful within-subject analysis makes false positives unlikely.

[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 (non-neighboring) village for the other decision. For this paper, we only use the same village pairs. The randomization used to pair participants was conducted within an experimental session.

[10] 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] Details on how this exclusion influences results can be found in Appendix E.

[12] These numbers are conditional on both participants in a pair having fully understood the instructions, which excludes about 8% of participants and their respective counterparts.

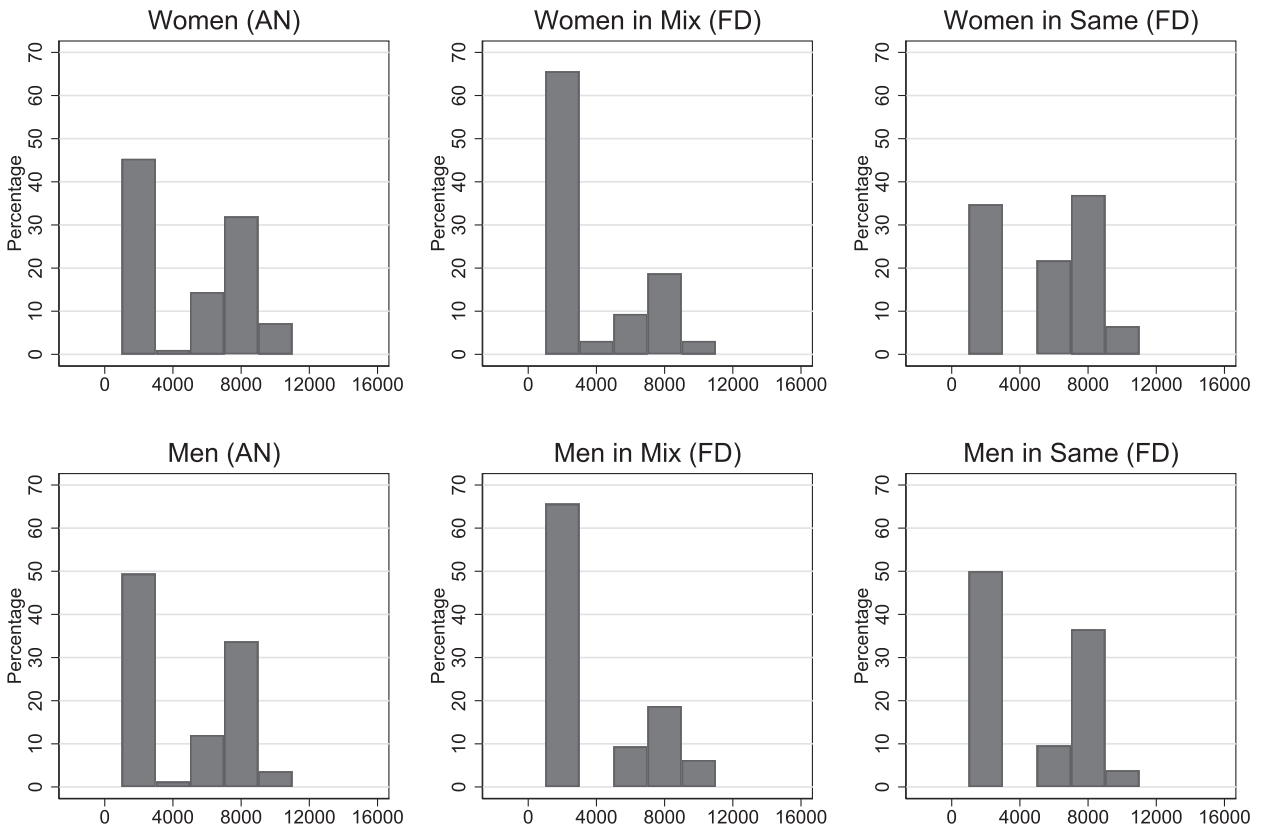

Fig. 1. Distribution of Individual Earnings, $N = 342$ (162 FD, 180 AN). Histograms show distribution of earnings by treatment and gender/ gender pairing; left-most bars also reflect frequency of disagreements.

3.1. Earnings

Figure 1 displays distributions of individual earnings by information condition, gender in AN, and gender pairing in FD. 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 and the panels in 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. In each panel, the left-most bar also 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 discernible. Perhaps most eye-catching are 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. This difference is economically meaningful as it amounts to almost 11% of the whole pie. In contrast, 'agreed earnings' are similar between genders. Indeed, across all gender pairings agreed earnings vary only between 7273 UGX and 7636 UGX.

To test whether there are statistically significant differences between genders, we conduct regression analyses with the following specification separately for AN and FD:[13]

$$y_{ij} = \beta_0 + \beta_1 \, \mathrm{G}_{ij} + \beta_2 \, \mathrm{X}_{ij} + \mu_k + \epsilon_{ij}, \tag{1}$$

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

[13] For detailed descriptive statistics, including non-parametric tests of gender differences in individual earnings, see Appendix Table C.1. The results reported there are in line with those of the presented regression analysis.

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

	Final Earnings		Agreed Earnings	
			AN	
	(1)	(2)	(3)	(4)
Men	-424.6	-365.1	-296.7	-453.2
	(480.6)	(547.3)	(369.9)	(370.9)
Constant	5889.4***	6264.0***	7735.1***	9220.5***
	(397.2)	(1773.6)	(231.5)	(1448.1)
Observations	180	179	96	95
R^2	0.063	0.113	0.014	0.263
			FD	
	(1)	(2)	(3)	(4)
WW	1728.4***	1648.2**	161.0	250.1
	(647.0)	(824.6)	(551.3)	(602.3)
MM	977.3	1300.8	310.8	320.2
	(1023.9)	(1141.0)	(537.6)	(631.4)
MW	125.0	242.4	363.6	650.9
	(256.7)	(233.4)	(786.5)	(770.9)
Constant	3799.5***	-205.8	7253.5***	5476.6***
	(571.3)	(5193.8)	(543.6)	(2059.3)
Observations	162	162	78	78
R^2	0.056	0.133	0.008	0.134
			Wald tests (p-values)	
WW - MM	0.324	0.755	0.285	0.853
MM - MW	0.396	0.372	0.887	0.529
WW - MW	0.009***	0.070*	0.544	0.318
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 village level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels. In AN, controls include risk aversion, age, education, wealth, trust, expected fairness, and expected generosity of ego. In FD, the same set of controls is included for ego's characteristics, plus controls for alter's characteristics and for the existence of a friendship tie between ego and alter. Coefficients of the controls are displayed in Appendix Tables D.1 and D.2.

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.[14] This vector includes wealth, age, education, trust, expected fairness, expected generosity, and risk aversion of ego (and alter for FD). Further, a control for the network variable indicating whether ego and alter are close friends is included in FD. The variable for wealth is an index that is obtained from a principal component analysis (see Appendix B for details). Age and education are measured in years. Expected fairness is measured on a scale from 1 (most negative) to 3 (most positive), while trust and expected generosity are measured on a reverse scale from 1 (most positive) to 3 (most negative). Risk aversion is measured on a four-point scale with a larger value reflecting higher risk aversion. μ_k captures experimenter fixed effects and ϵ_{ij} the remaining idiosyncratic error. We use bootstrapped cluster-robust standard errors with 2000 repetitions to adjust for potential non-independencies within a village.[15]

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. 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 formed by women in a mixed pair (WM). Wald tests on the bottom of the FD panel present comparisons between the gender pairings.

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 positive coefficient of WW shows that women in a same-gender pair have substantially higher final earnings than women in a mixed-gender pair. The 1648.2 UGX (Model 2 correspond to

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

[15] Here we follow Cameron et al. (2008) who advise to use bootstrapping if the number of clusters is low. Data from 14 sessions are included in the analysis, which draw on participants from 24 villages. See Appendix Tables A.2 and A.3 for details.

more than 10% of the resources at stake and almost 14% of the surplus (i.e., resources minus disagreement earnings). Further, women in same-gender pairs also earn substantially and significantly more than men in mixed-gender pairs, as can be seen from the difference in coefficients between WW and MW (1405.8 UGX or 11.7% of the surplus, Model 2) and the significant comparison reported in the panel below the regression table. On the other hand, for men, the gender of the counterpart does not influence final earnings. The coefficient of MM is not significant, and neither is the comparison between MM and MW 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 a first result.

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

The result that women tend to earn more when matched with another woman when identities are revealed together with the absence of gender differences in agreed earnings suggests that the observed differences in final earnings for pairs in FD are likely related to differences in disagreement rates between gender pairs. Next we analyze this possibility.

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 Expression (1), again separately for AN and FD.[16] As disagreement is an outcome on the pair level, in the analysis 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 for both AN and FD, mixed-gender pairs form the reference category, against which outcomes of WW and MM pairs are compared.

Table 2 reports the results of the regression analysis by gender pairing in AN and FD.[17] Model 1 shows results without controls, whereas Model 2 includes ego and alter controls for wealth, age, education, trust, expected fairness, expected generosity, and risk aversion, as well as a control for the existence of a friendship tie. In AN, no difference in disagreement between different gender pairings is observed. For FD, however, results are different. The significant and negative coefficient of WW in Model 1 shows that the probability of disagreement is almost 31 percentage points lower among WW pairs than among mixed-gender pairs. After introducing controls in Model 2, this effect is still substantial (29 percentage points) but at a weaker significance level. 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 about 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. Based on these observations, we now turn to an analysis of the possible mechanisms for the observed gender (pairing) differences.

4. Mechanisms

In this section we will explore some potential mechanisms that could explain the gender differences in earnings and disagreement. Specifically, we will test whether these differences in outcomes might be driven by gender differences in demands or beliefs. In addition, as gender might correlate with a variety of socio-economic characteristics, we will test whether any of these characteristics drive gender differences in demands or beliefs.

To gain insight into the latter aspect, Table 3 displays relevant socioeconomic characteristics separated by gender. The p-values in the last column indicate that there exists a gender difference for education and expected generosity. Specifically, women have lower education levels than men, and expect others to be slightly more generous. Risk aversion and friendship ties do not differ by gender.[18]

In the following analysis of the role of demands and beliefs in bargaining we will explore the effects of those characteristics that are significantly different between men and women (i.e., education and expected generosity). This way we

[16] Specifically, we employ a linear probability regression model. See Tables D.5 and D.6 in the appendix for results based on a logit regression. The results are qualitatively the same.

[17] For detailed descriptive statistics including non-parametric tests of gender differences in pairwise disagreement, see Appendix Table C.2. Results are in line with the presented regression analysis.

[18] Table 3 addresses the degree (number of friends) that men and women have on average. Appendix Table B.2 further shows that there is no difference in the frequency with which men and women are friends with participants from the same or the other gender.

Table 2
The Effect of Gender Pairings on Pairwise Disagreement.

	AN	
	(1)	(2)
WW	-0.117	-0.161
	(0.119)	(0.140)
MM	-0.0268	-0.180
	(0.152)	(0.191)
Constant	0.372***	0.509
	(0.0881)	(1.049)
Observations	90	89
R^2	0.069	0.259
	Wald tests (p-values)	
WW - MM	0.608	0.933
	FD	
	(1)	(2)
WW	-0.308***	-0.292*
	(0.0998)	(0.156)
MM	-0.156	-0.188
	(0.182)	(0.245)
Constant	0.658***	1.269
	(0.112)	(1.023)
Observations	81	81
R^2	0.064	0.213
	Wald tests (p-values)	
WW - MM	0.278	0.668
Controls	No	Yes

Linear regressions. Dependent variable is disagreement. Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at village level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels. Controls in AN and FD include risk aversion, age, education, wealth, trust, expected fairness, and expected generosity of ego and alter, plus a control for the existence of a friendship tie between ego and alter. Coefficients of the controls are displayed in the Appendix, Tables D.3 and D.4.

can gain a better understanding of the role these characteristics play in the observed gender differences in earnings and disagreements.[19]

Demands To test whether there are any gender differences with regard to demands, we run a regression model equivalent to the one described in Expression (1), using demands as dependent variable. Models 1 to 4 of Table 4 report the results, for AN in the upper panel and for FD in the lower panel. Model 1 shows results without expected generosity and education as controls, whereas Model 2 includes expected generosity, Model 3 includes education, and Model 4 both (of ego in AN, and of ego and alter in FD). As in the earnings regressions, women form the reference category in the AN regressions, while WM is the reference category in the FD regressions.

In AN, no significant gender differences are found, indicating that gender does not matter for demands if the identity of the counterpart is not known. In FD, however, differences between genders are observed. The significant difference between WW and MW pairs (reported in the panel below the table) indicates that men who are paired with a woman demand more than a woman who is paired with another woman. Women's demand in mixed-gender pairs do not differ from their demands in same-gender pairs, as indicated by the insignificant coefficient of WW. Further, the coefficient of MM is not statistically significant, which indicates that, if paired with a man, men and women make similar demands. All of these results are robust to the inclusion of education and expected generosity as controls and these controls are not significant themselves.[20] 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.

[19] Detailed descriptive statistics of demands, beliefs, and the combination thereof, including non-parametric tests for possible gender differences, are presented in Appendix Tables C.3, C.4, and C.5. Results are in line with the presented regression analysis.

[20] As would be expected based on Table 3, results are also robust to the inclusion of the other socioeconomic controls.

Table 3
Gender and Socioeconomic Characteristics.

	Women			Men			
	N	mean	st.dev.	N	mean	st.dev.	p-value[+]
Risk aversion	106	2.028	0.822	93	1.882	0.845	0.146
Age	106	43.311	14.611	93	39.914	13.709	0.100
Education	106	5.047	3.522	92	6.804	3.336	<0.001
Wealth	106	-0.358	1.913	93	0.078	2.190	0.108
Trust	106	2.113	0.421	93	2.172	0.433	0.381
Expected Fairness	106	2.047	0.541	93	2.140	0.653	0.230
Expected Generosity	106	1.906	0.669	93	2.161	0.696	0.009
Friend	106	6.094	3.961	93	6.796	4.742	0.413

[+] Two-sided p-value of a Mann-Whitney U test. Risk aversion is a self-reported measure on a scale from 1 (low risk aversion) to 4 (high risk aversion). Education measures years of schooling. Wealth is the first component of a principal component analyses (calculated on the full survey sample). Expected Fairness is measured on a scale from 1 (most negative) to 3 (most positive). Trust and Expected Generosity are measured in reverse order, on a scale from 1 (most positive) to 3 (most negative). Friend denotes the amount of people in the community that a respondent calls a close friend.

Table 4
The Effect of Gender (Pairings) on Demands and Beliefs.

	Demands		Beliefs					
				AN				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male	127.1	163.8	-19.22	-7.660	173.0	90.66	129.7	31.55
	(320.1)	(341.0)	(321.8)	(339.2)	(275.7)	(306.8)	(294.6)	(331.0)
Exp. Generosity		-112.2		-30.83		251.6		261.9
		(288.9)		(283.5)		(285.7)		(298.1)
Education			90.44*	89.38*			14.76	23.82
			(48.73)	(49.15)			(39.60)	(40.31)
Constant	8223.0***	8431.0***	7662.8***	7726.7***	6887.3***	6420.8***	6806.6***	6263.8***
	(345.8)	(672.4)	(433.5)	(747.2)	(328.8)	(647.5)	(410.6)	(764.4)
Observations	180	180	179	179	180	180	179	179
R^2	0.021	0.022	0.038	0.038	0.053	0.059	0.053	0.059
				FD				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WW	-330.1	-299.6	-275.9	-230.3	231.3	165.3	144.6	52.15
	(350.9)	(350.1)	(311.6)	(306.3)	(543.2)	(539.7)	(545.0)	(529.6)
MM	377.1	374.3	439.3	439.4	81.55	60.01	-251.6	-289.7
	(610.6)	(601.1)	(652.2)	(641.2)	(536.5)	(520.8)	(576.9)	(562.7)
MW	406.3	472.6	513.8	611.8*	-156.3	-295.5	-298.5	-491.7*
	(373.1)	(374.4)	(347.5)	(358.3)	(372.8)	(332.3)	(345.9)	(284.3)
Exp. Generosity ego		-4.109		-30.37		95.30		148.5
		(185.0)		(189.1)		(181.1)		(189.3)
Exp. Generosity alter		231.9		258.7		-399.9*		-421.5*
		(213.2)		(214.7)		(238.2)		(230.2)
Education ego			-30.17	-32.34			84.32*	90.36**
			(41.44)	(42.88)			(44.14)	(42.38)
Education alter			27.20	33.92			8.455	-3.029
			(51.96)	(54.25)			(61.36)	(59.94)
Constant	8313.3***	7825.8***	8276.6***	7745.9***	7147.5***	7824.8***	6682.9***	7355.6***
	(546.4)	(759.9)	(701.3)	(946.3)	(512.3)	(708.5)	(677.1)	(980.4)
Observations	162	162	162	162	162	162	162	162
R^2	0.028	0.034	0.033	0.040	0.027	0.046	0.045	0.067
				Wald tests (p-values)				
WW - MM	0.107	0.128	0.168	0.077*	0.790	0.848	0.476	0.422
MM - MW	0.958	0.862	0.908	0.944	0.623	0.470	0.933	0.883
WW - MW	0.006***	0.006***	0.008***	0.006***	0.374	0.297	0.299	0.250

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

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. Thus, men's higher demands in mixed pairs cause disagreements which leave not only themselves, but also those women they are matched with, worse off.

Beliefs Beliefs about the counterpart's demand are important determinants of outcomes in the NDG, as pointed out already by Harsanyi (1962). Therefore, we next explore whether beliefs play a role in the observed gender differences in bargaining outcomes. That is, we aim to understand if beliefs themselves depend on the gender of the opponent. In our regression analysis of beliefs, we again use Expression (1), now with beliefs as dependent variable. As before, women form the reference category in AN and WM is the reference category in FD. The results are reported in Models 5 to 8 of Table 4. As we only see a weak effect for MW in Model 8, we conclude that gender (and gender pairings in FD) do not robustly influence beliefs in either AN or FD. We state this observation in the following result.

Result 4. There are no gender differences in beliefs for either AN or 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 earnings and disagreement could be explained by potential gender differences in the relation between beliefs and demands. We look into two different concepts that link beliefs and demands: (1) 'belief inaccuracy', meaning that the expected demand of the other does not coincide with their actual demand, and (2) 'demand inconsistency', meaning that own demands and own beliefs do not add up exactly to the available resource.[21]

Demand Inconsistency Demand inconsistency is defined as the available resource of 16,000 - demand - belief, such that a larger value describes a larger demand inconsistency. To analyze whether gender differences in inconsistency exist, we run a regression based on Expression (1), with inconsistency as dependent variable. As before, women form the reference category in AN, and WM is the reference category in FD. Models 1 to 4 in Table 5 display results, for AN in the upper panel and for FD in the lower panel. As before, Model 1 shows results without controls, Model 2 includes expected generosity, Model 3 includes education and Model 4 include both controls. In AN the controls are included for ego only, and for FD the controls are for both ego and alter. In the latter case we also control for whether the bargaining partners have a friendship tie.

Results show that gender does not influence inconsistency in AN. Interestingly and intuitively, demand inconsistency decreases with more education in Models 3 and 4. In contrast, the FD regressions show some gender effects on demand inconsistency. The significantly negative coefficient of MM in Model 1 indicates that women in mixed pairs show a larger inconsistency in demands than men in same-gender pairs. Moreover, the marginally significant p-value of the comparison between WW and MM (reported below the FD panel) indicates that women in same-gender pairs tend to exhibit larger inconsistency in demands than men in same-gender pairs. Thus, irrespective of whether women are matched with a man or a woman, they display a larger demand inconsistency than men do in same-gender pairs. These results become somewhat weaker in Model 2, when expected generosity is included and they disappear entirely in Models 3 and 4, where education is included. We summarize these observations in the following result.

Result 5. (A) In AN, demand inconsistency of men and women does not differ. (B) In FD, irrespective of the gender of their counterpart, demands of women are more inconsistent than those of men. This result becomes weaker when controlling for expected generosity and disappears when controlling for education.

We conclude that both controls, expected generosity and education, appear to play a role in determining the gender differences in inconsistency, with education being the strongest factor. Importantly, women's higher demand inconsistency in FD apparently enables them to avoid disagreement when bargaining with other women. It benefits women in same-gender pairs because it gives a larger margin for agreement, whereas men's more consistent demands appear too aggressive, often leading to disagreement.

Belief Inaccuracy Next we analyze the role of belief inaccuracy, which is defined as the difference between a bargainer's belief and the actual demand of the opponent. To analyze gender differences in inaccuracy, we follow the same outline as for inconsistency. The results reported in Models 5 to 8 of Table 5 show that gender does not influence belief inaccuracy in either AN or FD. While inaccuracy is high for both men and women in AN and FD,[22] it does not differ between genders and gender pairings.

Result 6. Belief accuracy does not differ by gender in either AN or gender pairing in FD.

In sum, the analysis of possible mechanisms of gender differences in bargaining outcomes indicates that demands are the strongest driver. The observed gender difference in demands is strong, and it is robust to the inclusion of socioeconomic

[21] 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, 2009a; Eckel and Grossman, 2008), yet we do not see this in our data. One might speculate that this is due to the fact that we used a non-incentivized survey measure. However, Falk et al. (2018) also use a survey measure and do find higher risk aversion for women. In any case, we cannot exclude the possibility that a stronger desire to avoid risk increases the inconsistency of demands.

[22] For details see, Appendix Table C.5.

Table 5
The Effect of Gender (Pairings) on Inconsistency and Inaccuracy.

	Inconsistency				Inaccuracy			
					AN			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male	-300.1	-254.5	-110.5	-23.89	91.27	169.1	33.91	125.4
	(301.8)	(290.3)	(297.4)	(299.6)	(480.9)	(547.1)	(498.5)	(580.7)
Exp. Generosity		-139.4		-231.1		-237.9		-244.1
		(270.2)		(279.8)		(352.5)		(371.8)
Education			-105.2**	-113.2**			20.61	12.17
			(51.74)	(56.71)			(42.93)	(47.77)
Constant	889.7***	1148.1*	1530.6***	2009.6**	-1360.7**	-919.6	-1475.1**	-969.2
	(301.7)	(622.2)	(470.3)	(869.8)	(563.8)	(813.2)	(690.4)	(1050.3)
Observations	180	180	179	179	180	180	179	179
R^2	0.015	0.017	0.041	0.046	0.077	0.080	0.077	0.080
					FD			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WW	98.82	134.4	131.2	178.2	967.7	937.5	934.3	894.3
	(340.0)	(335.1)	(391.1)	(380.4)	(704.0)	(710.5)	(737.4)	(745.6)
MM	-458.7**	-434.3*	-187.7	-149.6	110.7	158.3	-177.0	-117.3
	(214.5)	(224.5)	(357.0)	(370.6)	(1029.2)	(1026.3)	(1158.1)	(1176.6)
MW	-250.0	-177.1	-215.3	-120.1	250.0	177.1	215.3	120.1
	(223.2)	(214.1)	(285.2)	(279.7)	(223.2)	(214.1)	(285.2)	(279.7)
Exp. Generosity ego		-91.19		-118.1		-136.6		-110.2
		(135.3)		(141.6)		(346.0)		(367.4)
Exp. Generosity alter		168.0		162.9		-395.8		-391.2
		(125.8)		(133.9)		(355.6)		(350.6)
Education ego			-54.15	-58.02			57.11	56.44
			(44.82)	(45.03)			(62.36)	(67.00)
Education alter			-35.66	-30.89			38.62	29.31
			(22.88)	(20.86)			(77.63)	(78.27)
Constant	539.2**	349.5	1040.5***	898.5**	-1572.0**	-473.6	-2107.5*	-1002.1
	(238.8)	(342.1)	(258.5)	(366.9)	(799.9)	(1337.1)	(1190.2)	(1750.8)
Observations	162	162	162	162	162	162	162	162
R^2	0.054	0.064	0.075	0.086	0.034	0.042	0.038	0.046
					Wald tests (p-values)			
WW - MM	0.062*	0.064*	0.296	0.175	0.373	0.412	0.281	0.293
MM - MW	0.292	0.190	0.914	0.661	0.891	0.985	0.727	0.823
WW - MW	0.242	0.324	0.202	0.323	0.276	0.260	0.257	0.228

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

controls.[23] Beliefs on their own do not play a role in inducing gender differences in bargaining outcomes, nor does belief inaccuracy. There are, however, discernible gender differences in demand inconsistency. Hence, in linking beliefs and demands, we observe a gender difference in inconsistency but not in inaccuracy. As inconsistency measures whether, given one's belief, all of the remaining resource is demanded, it can be interpreted as reflecting caution. Women leaving more of the resource undemanded show more cautious behavior than men. Specifically, inconsistency may be interpreted as reflecting awareness about own inability to estimate beliefs precisely. This awareness helps them to avoid disagreement in WW pairs.

5. Conclusion

In this study we present new insights from a controlled bargaining experiment in the field on the effect of gender and gender pairing on bargaining outcomes as well as potential mechanisms driving these 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-only pairs), two women (women-only pairs), or one man and one woman (mixed-gender pairs). To analyze the effect of gender and gender pairing on bargaining behavior, we exogenously varied identity disclosure. Depending on the information

[23] While the tables presented in this section show this for education and expected generosity, Appendix Tables D.7 show that this also holds true when including all socioeconomic controls.

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 earn less than women when the bargaining counterpart is a woman compared to when it is a men. Otherwise, men's final earnings do not depend on the gender of their counterpart. 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 about 30 percentage points lower among women-only pairs than among mixed-gender pairs, while disagreement is equally likely in men-only pairs and mixed-gender pairs. The lower disagreement rate of women-only pairs is in line with their higher final earnings, indicating that it is indeed this lower disagreement rate that enables women in same-gender pairs to secure higher earnings than those in mixed-gender pairs.

To better understand the observed gender differences in bargaining outcomes we analyzed whether gender differences in underlying potential mechanisms—demands, beliefs, and the relation between the two—could explain the results. Regarding demands, we found that, in FD, men demand more than women, but only when they are paired with a woman. This is in line with the observed gender differences in disagreement rates. 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.

We further saw that beliefs themselves do not differ across gender and gender pairings. Given that demands do differ, this suggests that gender differences in disagreement and earnings could be explained by potential gender differences in the relation between beliefs and demands. We introduced the concepts of belief inaccuracy and demand inconsistency to analyze this relation. A belief is inaccurate if it does not match the opponent's demand and is more inaccurate the larger the deviation is. A demand is inconsistent if one's own belief and demand does not add up to exactly the available resource; again larger deviations represent larger inconsistencies. Our results show that, in FD, irrespective of the gender of their counterpart, women exhibit higher demand inconsistencies than men, while belief inaccuracy does not differ across gender pairings. In women-only pairs these higher inconsistencies decrease the likelihood of disagreements and therefore benefits women in such pairs. However, larger inconsistencies comprise the possibility that in agreements more money is left on the table which could put such bargainers at a disadvantage in terms of agreed earnings. However, we do not find gender pairing differences in agreed earnings which suggests that demand inconsistencies are not large enough to affect women negatively in bargaining.

In our sample, we found gender differences in education and a measure of expected generosity and analyzed whether these characteristics may affect the considered mechanisms. When including these individual characteristics into the analysis of possible mechanisms, we found that the gender difference in demands is robust, whereas this is not the case for demand inconsistency. The latter gender effect becomes weaker when including expected generosity and disappears when including education. Hence, it appears that women's lower education and—to some extend—their more positive expectation regarding other's generosity are possible drivers of the observed gender difference in demand inconsistency.

An important conundrum this paper aimed to shed light on are observations indicating that women tend to perform better in negotiations in societies with lower levels of individualism, which also tend to be the societies with higher levels of gender inequality. This is also confirmed by our findings. Women obtained higher earnings than men when paired with another woman, while men obtained lower earnings independent of the gender of their opponent. How can this be reconciled with existing gender inequality? Importantly, women did not fare better in mixed gender pairs. In fact, their earnings were lowest in mixed gender pairs, albeit mostly insignificantly so. Given women's disadvantaged access to resources and opportunities, it is the mixed-gender pairs that are most important for women to increase their access to resources and opportunities. Development policy might have focused too much on supporting women groups in an attempt to increase women's empowerment (e.g. see the review in Brody et al., 2015) and should consider shifting its focus on making mixed-gender interactions more beneficial for women. Creating evidence on this conjecture may be an interesting avenue of future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Details on treatments and sampling

In the overall experiment, two different versions of the NDG were implemented, one where the disagreement payoff was symmetric and one where it was asymmetric. These versions where randomized between sessions. Only the NDG with symmetric disagreement payoff is used here (Table A.1).

For the NDG with symmetric disagreement payoff, four information conditions were implemented. One where participants were paired with somebody from a different village (DV), one where participants were paired with somebody from the same village but identities were not disclosed (anonymous – AN), one where participants were paired with somebody from the same village and one identity was disclosed while the other was not (semi-disclosure – SD), and one where participants were paired with somebody from the same village and both identities were disclosed (full disclosure – FD).

DV, AN, and FD were implemented as within subject. Also DV, AN, and SD were implemented as within subject. SD and FD, on the contrary were organized between subjects, with randomization taking place at the session level. That means, in an FD session, participants played DV, AN, and twice FD. In an SD session, on the contrary, participants played DV, AN, and twice SD. The order of the disclosure conditions was randomized at session level.

Table A.2 provides information on the selection of participants for this paper relative to all participants and Tables A.3 and A.4 show some more details on the sessions used, order of information conditions and the order of decisions made.

Table A.1
Treatments.

Symmetric	DV	**AN**	SD **FD**
Asymmetric	DV	AN	SD FD

Treatments implemented. Only Symmetric AN and Symmetric FD used here.

Table A.2
Sample Selection.

Village	All	This paper	Understood
1111	19	10	10
1112	16	9	9
1121	19	6	6
1131	18	16	16
1141	11	0	0
1142	17	17	16
1151	9	8	6
1152	21	0	0
1161	12	0	0
1412	32	16	14
1432	19	10	6
1511	14	7	5
1521	21	10	8
1522	21	11	9
1531	16	8	7
1542	16	8	4
1561	21	10	8
1562	23	12	11
1563	11	6	6
1621	24	4	4
1622	11	6	4
1623	10	0	0
1624	21	10	10
1631	20	20	14
1642	10	4	4
1721	16	6	6
1742	21	12	12
1751	21	7	4
Total	490	233	199

Column 'All' includes everybody who participated in the experiment. Column 'This paper' includes participants in Symmetric AN and Symmetric FD. Column 'Understood' includes only those participants from the column 'This paper' who understood the instructions (this is the final sample we included in the analysis).

Table A.3
Detailed Session Information.

Session	Information Conditions	Villages	N
3	AN, DV, SD, SD	1121, 1131	12
4	DV, FD, FD, AN	1131, 1151	16
6	AN, FD, FD, DV	1112, 1142	16
8	FD, DV, AN, FD	1111, 1142	19
10	FD, FD, AN, DV	1531, 1562	18
11	DV, SD, SD, AN	1561, 1563	14
13	AN, DV, SD, SD	1521, 1542	12
16	AN, DV, FD, FD	1511, 1522	14
17	SD, AN, DV, SD	1412, 1432	20
19	AN, SD, SD, DV	1622, 1631	10
21	DV, SD, SD, AN	1621, 1631	12
23	AN, DV, SD, SD	1624, 1642	14
26	AN, DV, FD, FD	1742, 1751	12
27	SD, AN, DV, SD	1721, 1742	10
Total			199

The reported experiment was part of a larger study. The missing session numbers indicate sessions with different tasks that are not analyzed in this paper. Column 'Information Conditions' displays the order in which information conditions were presented. Column 'Villages' shows which villages a session drew on (two per session). N is the number of participants.

Table A.4
Order of Decisions.

Order / Demand	AN			FD		
AN first	137	8386.86	2197.08	95	8631.58	1907.73
FD first	43	9325.58	2661.19	67	8738.81	2073.20
p-value		0.041			0.926	

Table displays average demands in AN and FD, depending on which of the information conditions was played first. To simplify, we only distinguish two cases. First row shows all decisions where AN was taken first. Second row shows all decisions where FD was taken first. Note that all AN decisions that were taken in an SD session are part of the first row.

Appendix B. Socioeconomic Characteristics

Table B.1 gives an overview of the gender composition and socioeconomic characteristics of participants.

Calculation of 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:[24]

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)

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 3 in the main text and Table B.1.

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

Table B.1
Gender and Socioeconomic Characteristics (depending on gender of alter).

Alter is a man	Women			Men			
	N	mean	st.dev.	N	mean	st.dev.	p-value$^+$
Risk aversion	77	2.078	0.791	90	1.933	0.845	0.180
Age	77	44.364	14.599	90	38.644	12.943	0.009
Education	77	4.338	3.485	90	7.544	3.332	<0.001
Wealth	77	-0.328	1.656	90	-0.063	2.329	0.711
Trust	77	2.117	0.396	90	2.133	0.342	0.941
Exp. Fairness	77	1.987	0.525	90	2.067	0.632	0.325
Exp. Generosity	77	1.870	0.676	90	2.156	0.669	0.008
Friend	77	6.338	3.868	90	6.344	4.717	0.777
Alter is a woman	Women			Men			
	N	mean	st.dev.	N	mean	st.dev.	p-value$^+$
Risk aversion	98	2.092	0.863	77	1.727	0.621	0.004
Age	98	43.969	14.329	77	39.442	12.935	0.030
Education	98	5.102	3.492	77	6.224	2.996	0.021
Wealth	98	-0.529	1.753	77	-0.074	2.155	0.131
Trust	98	2.112	0.428	77	2.156	0.431	0.568
Exp. Fairness	98	2.031	0.546	77	2.182	0.623	0.074
Exp. Generosity	98	1.918	0.668	77	2.169	0.657	0.016
Friend	98	5.929	3.859	77	6.117	4.344	0.978

Two-sided p-value of a Mann-Whitney U test. Upper panel denotes observations where alter is a man, lower panel where alter is a woman. 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 is the first component of a principal component analyses (calculated on the full survey sample). Exp. Fairness is measured on a scale from 1 (most negative) to 3 (most positive). Trust and Exp. Generosity are measured in reverse order, on a scale from 1 (most positive) to 3 (most negative). Friend denotes the number of people in the community that a respondent calls a close friend.

Table B.2
Pairwise Friendship Ties.

	WW	MM	Mix
N	49	45	77
%	34.69	35.56	36.36

Second row denotes the percentage of pairs that have a friendship tie. Chi-square test between the different gender pairings: p-value = 0.982.

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 aversion in Table 3 and Table B.1 reports descriptive statistics and Fig. B.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 ($N = 199$). 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.

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 suggests that the above observation, 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.

Table B.2 shows the frequency of friendship ties separated by gender pairings. It can be seen that these frequencies are very similar across gender pairings.

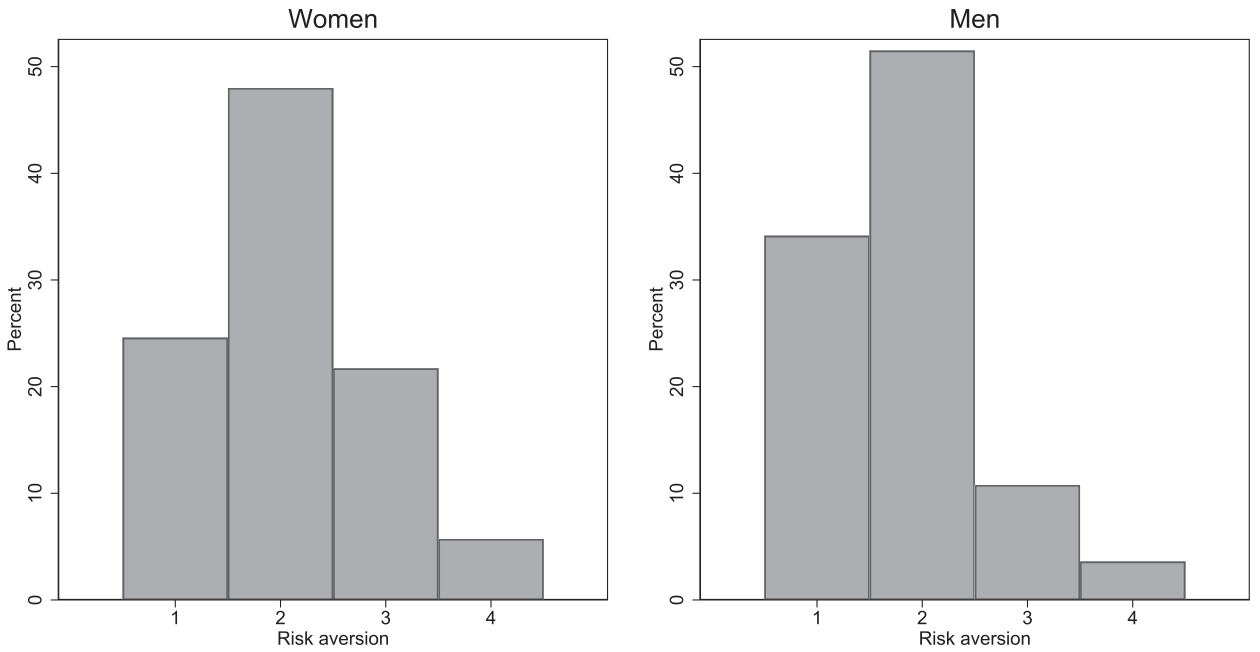

Fig. B1. Risk Aversion.

Appendix C. Non-parametric tests

Tables C.1 to C.5 report results of non-parametric tests.

Table C.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 FD 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. Numbers in the bottom right corners of the panels indicate two-sided p-values of a Mann-Whitney U test between WW and MM.

Table C.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 C.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 FD 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 C.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 FD 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 C.5
Belief Inccuracy and Demand Inconsistency.

	Woman			Man			
	N	mean	st.dev.	N	mean	st.dev.	*p*-value[+]
a) Belief Inaccuracy							
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 Inconsistency							
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 FD 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. Numbers in the bottom right corners of each panel indicate two-sided *p*-values of a Mann-Whitney *U* test between WW and MM.

Appendix D. Additional regression tables

Tables D.1 to D.8 present additional regression tables.

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

	Final Earnings		Agreed Earnings	
	(1)	(2)	(3)	(4)
Men	-424.6	-365.1	-296.7	-453.2
	(480.6)	(547.3)	(369.9)	(370.9)
Risk aversion ego		-286.7		-810.1***
		(200.8)		(231.3)
Age ego		-5.409		13.78
		(16.75)		(12.88)
Education ego		-46.11		76.62
		(86.38)		(69.27)
Wealth ego		171.7		-14.84
		(134.3)		(92.07)
Trust ego		222.1		-63.59
		(520.0)		(379.4)
Exp. Fairness ego		813.2*		375.4
		(467.1)		(331.8)
Exp. Generosity ego		-746.0*		-823.8**
		(396.5)		(382.5)
Constant	5889.4***	6264.0***	7735.1***	9220.5***
	(397.2)	(1773.6)	(231.5)	(1448.1)
Observations	180	179	96	95
R^2	0.063	0.113	0.014	0.263

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

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

	Final Earnings		Agreed Earnings	
	(1)	(2)	(3)	(4)
WW	1728.4***	1648.2**	161.0	250.1
	(647.0)	(824.6)	(551.3)	(602.3)
MM	977.3	1300.8	310.8	320.2
	(1023.9)	(1141.0)	(537.6)	(631.4)
MW	125.0	242.4	363.6	650.9
	(256.7)	(233.4)	(786.5)	(770.9)
Risk aversion ego		77.82		-63.49
		(364.8)		(360.6)
Age ego		28.32*		12.32
		(16.28)		(19.27)
Education ego		-43.34		-47.17
		(103.8)		(98.41)
Wealth ego		128.3		-0.645
		(135.6)		(125.7)
Trust ego		779.1		-527.1
		(898.0)		(627.9)
Exp. Fairness ego		-8.977		131.4
		(529.3)		(399.1)
Exp. Generosity ego		-479.9		227.8
		(455.6)		(339.3)
Friend		-313.5		-111.6
		(667.0)		(342.2)
Risk aversion alter		155.3		141.4
		(359.6)		(278.0)
Age alter		22.74		7.287
		(22.82)		(21.87)

(continued on next page)

Table D.2
(continued)

	Final Earnings		Agreed Earnings	
	(1)	(2)	(3)	(4)
Education alter		24.11		110.3
		(105.5)		(105.5)
Wealth alter		155.9		44.69
		(109.9)		(122.4)
Trust alter		1016.0		256.7
		(1156.2)		(521.7)
Exp. Fairness alter		-97.62		91.17
		(618.5)		(272.5)
Exp. Generosity alter		-464.5		43.41
		(500.8)		(265.5)
Constant	3799.5***	-205.8	7253.5***	5476.6***
	(571.3)	(5193.8)	(543.6)	(2059.3)
Observations	162	162	78	78
R^2	0.056	0.133	0.008	0.134
			Wald tests (p-values)	
WW - MM	0.324	0.755	0.285	0.853
MM - MW	0.396	0.372	0.887	0.529
WW - MW	0.009***	0.070**	0.544	0.318

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 village level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

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

	Disagreement	
	(1)	(2)
WW	-0.117	-0.161
	(0.119)	(0.140)
MM	-0.0268	-0.180
	(0.152)	(0.191)
Risk aversion ego		-0.0596
		(0.0659)
Age ego		0.00522
		(0.00568)
Education ego		0.00421
		(0.0294)
Wealth ego		-0.0293
		(0.0385)
Trust ego		-0.129
		(0.195)
Exp. Fairness ego		-0.158
		(0.119)
Exp. Generosity ego		0.0905
		(0.0988)
Friend		-0.256**
		(0.127)
Risk aversion alter		0.0803
		(0.0898)
Age alter		0.00239
		(0.00463)
Education alter		0.0341
		(0.0219)
Wealth alter		-0.0419
		(0.0355)

(continued on next page)

Table D.3
(continued)

	Disagreement	
	(1)	(2)
Trust alter		-0.0202
		(0.171)
Exp. Fairness alter		-0.0868
		(0.149)
Exp. Generosity alter		0.0243
		(0.105)
Constant	0.372***	0.509
	(0.0881)	(1.049)
Observations	90	89
R^2	0.069	0.259
Wald tests (p-values)		
WW - MM	0.608	0.933

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

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

	Disagreement	
	(1)	(2)
WW	-0.308***	-0.292*
	(0.0998)	(0.156)
MM	-0.156	-0.188
	(0.182)	(0.245)
Risk aversion ego		-0.0273
		(0.108)
Age ego		-0.00640
		(0.00853)
Education ego		-0.00916
		(0.0279)
Wealth ego		-0.0334
		(0.0395)
Trust ego		-0.178
		(0.311)
Exp. Fairness ego		0.150
		(0.167)
Exp. Generosity ego		0.105
		(0.181)
Friend		0.0384
		(0.127)
Risk aversion alter		-0.0296
		(0.0957)
Age alter		-0.00408
		(0.00548)
Education alter		0.0211
		(0.0367)
Wealth alter		0.0113
		(0.0393)
Trust alter		-0.104
		(0.255)
Exp. Fairness alter		-0.113
		(0.114)
Exp. Generosity alter		0.0754
		(0.134)
Constant	0.658***	1.269
	(0.112)	(1.023)

(continued on next page)

Table D.4
(continued)

	Disagreement	
	(1)	(2)
Observations	81	81
R^2	0.064	0.213
Wald tests (p-values)		
WW - MM	0.278	0.668

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

Table D.5
The Effect of Gender Pairings on Pairwise Disagreement in AN (Logit).

	Disagreement	
	(1)	(2)
WW	-0.504	-0.785
	(0.559)	(1077.0)
MM	-0.116	-1.151
	(0.729)	(731.0)
Risk aversion ego		-0.406
		(732.9)
Age ego		0.0317
		(21.52)
Education ego		-0.000740
		(126.0)
Wealth ego		-0.141
		(158.8)
Trust ego		-0.666
		(1499.4)
Exp. Fairness ego		-0.940
		(711.2)
Exp. Generosity ego		0.530
		(1287.4)
Friend		-1.357
		(1372.9)
Risk aversion alter		0.482
		(1221.7)
Age alter		0.0152
		(44.06)
Education alter		0.217
		(300.0)
Wealth alter		-0.257
		(393.1)
Trust alter		-0.113
		(507.3)
Exp. Fairness alter		-0.539
		(1057.6)
Exp. Generosity alter		0.234
		(1334.3)
Constant	-0.528	-0.230
	(0.411)	(9481.6)
Observations	90	89
R^2		
Wald tests (p-values)		
WW - MM	0.644	0.999

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

Table D.6
The Effect of Gender Pairings on Pairwise Disagreement in FD (Logit)

	Disagreement	
	(1)	(2)
WW	-1.274***	-1.402
	(0.463)	(5058.3)
MM	-0.647	-0.925
	(0.847)	(8962.9)
Risk aversion ego		-0.136
		(3223.4)
Age ego		-0.0298
		(291.8)
Education ego		-0.0413
		(1299.8)
Wealth ego		-0.176
		(321.2)
Trust ego		-0.836
		(14053.9)
Exp. Fairness ego		0.703
		(6358.5)
Exp. Generosity ego		0.506
		(5862.7)
Friend		0.179
		(5141.8)
Risk aversion alter		-0.202
		(3561.4)
Age alter		-0.0187
		(216.6)
Education alter		0.1000
		(1099.2)
Wealth alter		0.0518
		(961.0)
Trust alter		-0.562
		(5583.1)
Exp. Fairness alter		-0.605
		(2491.1)
Exp. Generosity alter		0.372
		(3684.3)
Constant	0.655	3.987
	(0.531)	(18924.8)
Observations	81	81
R^2		
Wald tests (p-values)		
WW - MM	0.335	0.999

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

Table D.7
Mechanisms including all controls - FD.

	Demands (1)	Beliefs (2)	Inconsistency (3)	Inaccuracy (4)
WW	-220.0	5.914	214.1	880.3
	(378.4)	(568.3)	(469.2)	(877.6)
MM	530.8	-317.8	-213.1	-194.3
	(713.0)	(606.8)	(443.7)	(1187.7)
MW	654.3	-618.8	-35.54	35.54
	(443.9)	(385.2)	(325.1)	(325.1)
Risk aversion ego	-43.24	133.1	-89.87	-21.31
	(221.4)	(231.4)	(162.8)	(340.8)
Age ego	4.951	-7.502	2.551	-9.218
	(19.21)	(16.45)	(9.131)	(20.87)
Educ ego	-15.63	49.90	-34.27	12.68
	(58.68)	(62.63)	(53.84)	(91.43)
Wealth ego	-114.4**	167.2*	-52.80	199.5*
	(51.98)	(87.47)	(77.76)	(120.0)
Trust ego	-248.0	-174.5	422.5	432.0
	(842.3)	(795.3)	(436.9)	(1137.4)
Exp. Fairness ego	-14.81	65.39	-50.58	-13.97
	(419.3)	(397.2)	(110.1)	(530.5)
Exp. Generosity ego	28.05	196.4	-224.5**	-176.7
	(289.2)	(279.5)	(104.8)	(402.5)
Friend	-38.59	-160.3	198.9	-121.7
	(400.5)	(302.9)	(225.7)	(673.3)
Risk aversion alter	154.4	-129.6	-24.85	-86.34
	(213.1)	(217.8)	(175.7)	(360.9)
Age alter	1.715	-1.913	0.198	-6.865
	(11.35)	(14.19)	(6.829)	(27.85)
Educ alter	37.22	-18.89	-18.33	-3.264
	(63.44)	(71.06)	(28.98)	(101.0)
Wealth alter	-32.38	57.53	-25.14	171.9*
	(77.55)	(71.30)	(41.51)	(91.97)
Trust alter	-606.5	592.0	14.49	840.0
	(477.5)	(504.0)	(308.6)	(1091.7)
Exp. Fairness alter	79.36	-214.2	134.8	-199.4
	(246.2)	(254.6)	(175.0)	(596.0)
Exp. Generosity alter	373.1*	-490.4**	117.3	-518.5
	(219.4)	(248.6)	(163.7)	(382.8)
Constant	8395.6***	7746.8**	-142.4	-1303.1
	(2768.9)	(3828.7)	(1885.2)	(6268.3)
Observations	162	162	162	162
R^2	0.067	0.117	0.117	0.083
	Wald tests (p-values)			
WW - MM	0.206	0.633	0.255	0.378
MM - MW	0.848	0.642	0.577	0.850
WW - MW	0.027**	0.264	0.418	0.253

Linear regressions. Dependent variables are demands (model 1), beliefs (model 2), inconsistency (model 3), and inaccuracy (model 4).Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at village level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance.

Table D.8
Mechanisms including all controls - AN.

	Demands	Beliefs	Inconsistency	Inccuracy
	(1)	(2)	(3)	(4)
gender_ego	662.1*	-366.0	-296.1	-495.4
	(367.0)	(326.2)	(203.0)	(621.4)
Risk aversion ego	-81.84	193.1	-111.3	56.19
	(205.7)	(181.9)	(119.2)	(237.1)
Age ego	5.358	-6.837	1.479	-8.357
	(18.33)	(16.41)	(9.644)	(19.54)
Educ ego	-15.51	51.63	-36.12	7.083
	(62.16)	(65.45)	(44.56)	(89.79)
Wealth ego	-116.7***	175.8**	-59.08	232.5**
	(43.97)	(86.84)	(74.05)	(105.5)
Trust ego	-190.0	-173.7	363.7	214.3
	(809.4)	(740.5)	(383.5)	(962.0)
Exp. Fairness ego	-50.24	115.1	-64.82	21.21
	(411.2)	(366.2)	(109.3)	(377.6)
Exp. Generosity ego	19.81	170.7	-190.5*	-147.2
	(268.8)	(254.0)	(103.0)	(386.4)
Constant	8558.2***	6837.1***	604.7	-919.9
	(2315.6)	(2201.0)	(970.7)	(2556.2)
Observations	162	162	162	162
R^2	0.041	0.079	0.091	0.045

Linear regressions. Dependent variables are demands (model 1), beliefs (model 2), inconsistency (model 3), and inaccuracy (model 4). Standard errors in parentheses, estimated with bootstrapping (with 2000 repetitions), clustered at village level: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ indicate two-sided significance levels.

Appendix E. Regression tables including those who did not understand

Tables E.1 to E.6 report regression results with all partcipants, inlcuding those who did not correctly answer the comprehension questions.

Table E.1
The Effect of Gender on Individual Earnings in AN (including those who did not Understand).

	Final Earnings		Agreed Earnings	
	(1)	(2)	(3)	(4)
Men	-418.0	-367.3	-388.1	-536.4
	(439.7)	(512.7)	(328.4)	(346.6)
Risk ego		-52.50		-483.4***
		(173.3)		(185.3)
Age ego		-4.006		11.07
		(16.65)		(12.26)
Educ ego		-40.45		48.28
		(71.35)		(66.45)
Wealth ego		204.0		37.25
		(126.6)		(77.74)
Trust ego		-187.7		-190.8
		(449.4)		(410.9)
Exp. Fairness ego		938.7***		531.9**
		(359.3)		(258.3)
Exp. Generosity ego		-650.7*		-819.5***
		(365.4)		(310.7)
Constant	5418.3***	5652.3***	7677.9***	8809.4***
	(453.3)	(1515.0)	(194.2)	(1292.3)
Observations	222	221	114	113
R^2	0.033	0.086	0.017	0.190

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

Table E.2
The Effect of Gender Pairings on Individual Earnings in FD (including those who did not Understand).

	Final Earnings		Agreed Earnings	
	(1)	(2)	(3)	(4)
WW	928.7	996.7	-47.52	194.1
	(606.4)	(660.4)	(464.2)	(517.5)
MM	706.2	738.9	153.2	213.8
	(1031.3)	(1150.9)	(476.4)	(476.4)
MW	87.50	263.8	218.8	590.7
	(290.2)	(253.0)	(719.0)	(797.1)
Risk ego		162.3		18.68
		(324.2)		(191.5)
Age ego		26.36*		13.86
		(14.97)		(14.01)
Educ ego		14.21		-13.80
		(94.43)		(61.18)
Wealth ego		198.4*		41.63
		(110.4)		(92.37)
Trust ego		323.7		-159.9
		(614.4)		(451.3)
Exp. Fairness ego		114.8		231.6
		(538.1)		(331.2)
Exp. Generosity ego		-403.7		-26.77
		(307.7)		(307.7)
Tie		-847.2		-106.2
		(525.9)		(209.0)
Risk alter		200.9		67.76
		(277.8)		(157.8)
Age alter		18.52		-0.638
		(18.60)		(14.51)
Educ alter		61.74		74.46
		(115.0)		(63.51)
Wealth alter		193.3**		13.26
		(82.11)		(77.69)
Trust alter		377.4		144.1
		(744.0)		(403.8)
Exp. Fairness alter		-33.45		44.83
		(611.4)		(220.6)
Exp. Generosity alter		-258.1		182.8
		(333.4)		(244.1)
Constant	3752.2***	648.4	7480.1***	5348.2***
	(510.7)	(3315.6)	(497.3)	(1213.9)
Observations	198	198	96	96
R^2	0.028	0.116	0.008	0.112
		Wald tests (p-values)		
WW - MM	0.757	0.806	0.123	0.941
MM - MW	0.521	0.678	0.839	0.412
WW - MW	0.166	0.234	0.376	0.268

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

Table E.3
The Effect of Gender Pairings on Pairwise Disagreement in AN (including those who did not Understand).

	Disagreement	
	(1)	(2)
WW	-0.124	-0.111
	(0.109)	(0.115)
MM	-0.0605	-0.0929
	(0.144)	(0.193)
Risk ego		-0.0506
		(0.0503)
Age ego		0.00779
		(0.00517)
Educ ego		0.00674
		(0.0254)
Wealth ego		-0.0187
		(0.0397)
Trust ego		-0.0768
		(0.152)
Exp. Fairness ego		-0.122
		(0.0926)
Exp. Generosity ego		0.0867
		(0.0793)
Tie		-0.145
		(0.112)
Risk alter		0.0122
		(0.0675)
Age alter		-0.00166
		(0.00389)
Educ alter		0.0160
		(0.0157)
Wealth alter		-0.0324
		(0.0300)
Trust alter		0.0535
		(0.133)
Exp. Fairness alter		-0.109
		(0.125)
Exp. Generosity alter		0.0347
		(0.0814)
Constant	0.457***	0.468
	(0.0937)	(0.802)
Observations	111	110
R^2	0.041	0.175
Wald tests (p-values)		
WW - MM	0.710	0.930

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

Table E.4
The Effect of Gender Pairings on Pairwise Disagreement in FD (including those who did not Understand).

	Disagreement	
	(1)	(2)
WW	-0.176*	-0.199*
	(0.103)	(0.111)
MM	-0.115	-0.117
	(0.177)	(0.238)
Risk ego		-0.0416
		(0.0807)
Age ego		-0.00249
		(0.00600)
Educ ego		-0.00730
		(0.0260)
Wealth ego		-0.0470
		(0.0293)
Trust ego		0.00902
		(0.174)
Exp. Fairness ego		0.0699
		(0.142)
Exp. Generosity ego		0.0870
		(0.110)
Tie		0.142
		(0.122)
Risk alter		-0.0231
		(0.0642)
Age alter		-0.00466
		(0.00488)
Educ alter		0.000617
		(0.0274)
Wealth alter		-0.00876
		(0.0314)
Trust alter		-0.140
		(0.183)
Exp. Fairness alter		-0.0846
		(0.0911)
Exp. Generosity alter		0.0579
		(0.109)
Constant	0.681***	1.100*
	(0.0921)	(0.603)
Observations	99	99
R^2	0.035	0.152
Wald tests (p-values)		
WW - MM	0.646	0.711

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

Table E.5
The Effect of Gender (Pairings) on Demands and Beliefs (including those who did not Understand).

	Demands		Beliefs					
				AN				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male	77.55	118.4	-98.58	-75.88	235.4	195.5	217.3	172.2
	(337.5)	(360.4)	(347.4)	(378.7)	(344.3)	(382.9)	(367.2)	(415.1)
Exp. Generosity		-149.2		-70.97		145.7		141.1
		(254.2)		(268.3)		(271.8)		(287.8)
Education			99.04**	96.94*			2.243	6.421
			(47.64)	(50.93)			(36.73)	(39.08)
Constant	8377.8***	8654.9***	7826.2***	7970.0***	6802.7***	6532.0***	6796.9***	6511.1***
	(283.3)	(582.9)	(322.1)	(742.2)	(255.1)	(516.5)	(281.5)	(637.8)
Observations	222	222	221	221	222	222	221	221
R^2	0.009	0.011	0.030	0.030	0.036	0.038	0.036	0.037
				FD				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WW	-73.61	-5.858	2.629	91.78	-309.7	-372.5	-354.1	-437.7
	(452.5)	(467.6)	(431.3)	(452.5)	(560.6)	(558.6)	(553.8)	(542.6)
MM	322.5	366.9	555.8	621.3	83.58	60.59	-408.6	-452.0
	(597.7)	(597.9)	(607.7)	(604.7)	(555.9)	(556.3)	(584.7)	(572.1)
MW	262.5	406.9	444.6	631.6	-12.50	-153.2	-185.7	-367.6
	(437.8)	(444.7)	(403.0)	(452.5)	(393.5)	(341.9)	(364.2)	(324.8)
Exp. Generosity ego		-167.1		-191.7		109.3		133.7
		(342.6)		(347.5)		(212.7)		(212.0)
Exp. Generosity alter		314.0		337.6*		-359.7*		-381.3*
		(193.7)		(193.0)		(207.3)		(203.6)
Education ego			-60.63	-65.94			107.0**	112.1**
			(55.59)	(56.14)			(51.36)	(49.77)
Education alter			8.730	14.18			41.02	35.67
			(46.99)	(47.49)			(50.24)	(49.43)
Constant	8598.2***	8234.6***	8787.6***	8403.9***	6829.7***	7396.9***	6116.5***	6700.0***
	(530.0)	(841.7)	(646.8)	(887.3)	(557.9)	(853.7)	(598.9)	(868.7)
Observations	198	198	198	198	198	198	198	198
R^2	0.008	0.020	0.018	0.033	0.008	0.023	0.038	0.055
				Wald tests (p-values)				
WW - MM	0.407	0.435	0.262	0.281	0.540	0.495	0.926	0.980
MM - MW	0.910	0.942	0.851	0.987	0.838	0.656	0.660	0.874
WW - MW	0.212	0.104	0.075*	0.039**	0.491	0.607	0.659	0.860

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

Table E.6
The Effect of Gender (Pairings) on Inconsistency and Inaccuracy (including those who did not Understand).

	Inconsistency				Inaccuracy			
					AN			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male	-313.0	-314.0	-118.7	-96.27	66.53	138.4	1.088	85.59
	(254.0)	(245.6)	(235.6)	(237.1)	(518.1)	(574.7)	(566.7)	(641.7)
Exp. Generosity		3.501		-70.13		-262.5		-264.2
		(222.7)		(234.5)		(329.0)		(358.3)
Education			-101.3**	-103.4**			25.58	17.75
			(43.02)	(46.62)			(41.35)	(49.20)
Constant	819.6***	813.1	1376.9***	1518.9**	-1531.6***	-1044.0	-1664.6***	-1129.3
	(226.6)	(494.7)	(368.3)	(695.2)	(479.3)	(703.5)	(474.9)	(902.9)
Observations	222	222	221	221	222	222	221	221
R^2	0.017	0.017	0.043	0.043	0.041	0.044	0.041	0.044
					FD			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WW	383.4	378.4	351.5	345.9	26.37	40.17	87.82	102.1
	(305.0)	(290.0)	(331.3)	(319.5)	(742.0)	(721.1)	(710.3)	(688.3)
MM	-406.1**	-427.5**	-147.2	-169.2	23.56	100.5	-519.7	-441.7
	(164.0)	(180.2)	(282.9)	(290.9)	(999.9)	(1028.2)	(1071.0)	(1085.7)
MW	-250.0	-253.6	-258.9	-264.0	250.0	253.6	258.9	264.0
	(181.3)	(204.2)	(244.7)	(279.4)	(181.3)	(204.2)	(244.7)	(279.4)
Exp. Generosity ego		57.84		58.02		-204.7		-203.9
		(231.3)		(236.2)		(330.2)		(336.4)
Exp. Generosity alter		45.71		43.68		-192.6		-189.6
		(135.4)		(143.4)		(449.3)		(457.8)
Education ego			-46.37	-46.17			98.27*	97.93*
			(35.48)	(36.25)			(57.51)	(58.96)
Education alter			-49.75**	-49.85**			101.7	101.6
			(23.29)	(24.57)			(75.93)	(79.83)
Constant	572.1**	368.5	1095.9***	896.1**	-2031.1**	-1244.5	-3115.8***	-2335.5
	(252.2)	(375.0)	(270.4)	(379.3)	(807.6)	(1719.3)	(1092.0)	(1816.7)
Observations	198	198	198	198	198	198	198	198
R^2	0.069	0.070	0.092	0.093	0.001	0.005	0.019	0.022
					Wald tests (p-values)			
WW - MM	0.006***	0.008***	0.088*	0.092*	0.998	0.955	0.560	0.597
MM - MW	0.343	0.304	0.560	0.680	0.818	0.879	0.463	0.515
WW - MW	0.021**	0.041**	0.018**	0.041**	0.757	0.780	0.806	0.828

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

Appendix F. 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 16,000 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 16,000 UGX) each will get the amount demanded.

However, if the total is more than 16,000 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 6000 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 6000 UGX.

Imagine now that person 1 demanded 11,000 UGX and person 2 demanded 7000 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 2000 UGX and person 2 would get 2000 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 5000 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 8000 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 9000 UGX. How much would person 1 get if he demanded 9000 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 9000 UGX. How much would person 1 get if he demanded 6000 UGX? (6,000 UGX). How much would person 2 get? (9,000 UGX)
3. Imagine that person 2 believes that person 1 will demand 6000 UGX. How much would person 2 get if he demanded 6000 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 6000 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 won't 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 won't 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 4000 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 8000 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 6000 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]

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jebo.2022.10.024.

References

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

Ado, A., Kurosaki, T., 2014. Motives for Sharing in Developing Countries: Experimental Evidence from Jakarta. PRIMCED Discussion Paper Series, no. 53. Institute of Economic Research, Hitotsubashi University, Kunitachi.

Altmann, S., Dohmen, T., Wibral, M., 2008. Do the reciprocal trust less? Econ. Lett. 99 (3), 454–457.

Andreoni, J., Petrie, R., 2008. Beauty, gender and stereotypes: evidence from laboratory experiments. J. Econ. Psychol. 29 (1), 73–93.

Armantier, O., Treich, N., 2013. Eliciting beliefs: proper scoring rules, incentives, stakes and hedging. Eur. Econ. Rev. 62, 17–40.

Azrieli, Y., Chambers, C.P., Healy, P.J., 2018. Incentives in experiments: a theoretical analysis. J. Polit. Economy 126 (4), 1472–1503.

Azrieli, Y., Chambers, C.P., Healy, P.J., 2020. Incentives in experiments with objective lotteries. Exp. Econ. 23, 1–29.

Babcock, L., Laschever, S., 2003. Women Don't Ask: Negotiation and the Gender divide. Princeton University Press, Princeton, NJ.

Baron-Cohen, S., O'riordan, M., Stone, V., Jones, R., Plaisted, K., 1999. Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. J. Autism. Dev. Disord. 29 (5), 407–418.

Binzel, C., Fehr, D., 2013. Giving and sorting among friends: evidence from a lab-in-the-field experiment. Econ. Lett. 121 (2), 214–217.

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

Bolton, G.E., Katok, E., 1995. An experimental test for gender differences in beneficent behavior. Econ. Lett. 48 (3–4), 287–292.

Bowles, H.R., Babcock, L., Lai, L., 2007. Social incentives for gender differences in the propensity to initiate negotiations: sometimes it does hurt to ask. Organ. Behav. Hum. Decis. Process. 103 (1), 84–103.

Brackett, M.A., Salovey, P., 2006. Measuring emotional intelligence with the Mayer-Salovery-Caruso emotional intelligence test (MSCEIT). Psicothema 18, 34–41.

Brody, C., De Hoop, T., Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., Dworkin, S.L., 2015. Economic self-help group programs for improving women's empowerment: a systematic review. Campbell Syst. Rev. 11 (1), 1–182.

Camerer, C.F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., Wu, H., 2016. Evaluating replicability of laboratory experiments in economics. Science 351 (6280), 1433–1436.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-based improvements for inference with clustered errors. Rev. Econ. Stat. 90 (3), 414–427.

Cardenas, J.-C., Dreber, A., von Essen, E., Ranehill, E., 2012. Gender differences in competitiveness and risk taking: comparing children in colombia and Sweden. J. Econ. Behav. Organ. 83 (1), 11–23.

Castillo, M., Petrie, R., Torero, M., Vesterlund, L., 2013. Gender differences in bargaining outcomes: a field experiment on discrimination. J. Public Econ. 99, 35–48.

Charness, G., Gneezy, U., 2012. Strong evidence for gender differences in risk taking. J. Econ. Behav. Organ. 83 (1), 50–58.

Cox, J.C., Deck, C.A., 2006. When are women more generous than men? Econ. Inq. 44 (4), 587–598.

Croson, R., Gneezy, U., 2009a. Gender differences in preferences. J. Econ. Lit. 47 (2), 448–474.

Croson, R., Gneezy, U., 2009b. Gender differences in preferences. J. Econ. Lit. 47 (2), 448–474.

D'Exelle, B., Riedl, A., 2019. Gender inequality and resource sharing: evidence from rural Nicaragua. J. Dev. Perspect. 2, 62–87.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2008. Representative trust and reciprocity: prevalence and determinants. Econ. Inq. 46 (1), 84–90.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2012. The intergenerational transmission of risk and trust attitudes. Rev. Econ. Stud. 79 (2), 645–677.

Dohmen, T.J., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2005. Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey. IZA Discussion Paper, No. 1730. Institute of Labor Economics, Bonn.

Eckel, C., Grossman, P., 2001. Chivalry and solidarity in ultimatum games. Econ. Inq. 39 (2), 171–188.

Eckel, C.C., Grossman, P.J., 1996. Altruism in anonymous dictator games. Games Econ. Behav. 16 (2), 181–191.

Eckel, C.C., Grossman, P.J., 1998. Are women less selfish than men? Evidence from dictator experiments. Econ. J. 108 (448), 726–735.

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

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

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences*. Q. J. Econ. 133 (4), 1645–1692.

Falk, A., Hermle, J., 2018. Relationship of gender differences in preferences to economic development and gender equality. Science 362, 307.

Gneezy, U., Leonard, K.L., 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., Rustichini, A., 2003. Performance in competitive environments: gender differences. Q. J. Econ. 118 (3), 1049–1074.

Gowdy, J., Iorgulescu, R., Onyeiwu, S., 2003. Fairness and retaliation in a rural Nigerian village. J. Econ. Behav. Organ. 52 (4), 469–479.

Harsanyi, J.C., 1962. Bargaining in ignorance of the opponent's utility function. J. Confl. Resolut. 6 (1), 29–38.

Hernandez-Arenaz, I., Iriberri, N., 2018. Women ask for less (only from men): evidence from bargaining in the field. J. Econ. Behav. Organ. 152, 192–214.

Hernandez-Arenaz, I., 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., Mackenbach, J.P., 2003. Measuring health inequality among children in developing countries: does the choice of the indicator of economic status matter? Int. J. Equity Health 2 (1), 8.
Jackson, C., 2013. Cooperative conflicts and gender relations: experimental evidence from Southeast Uganda. Fem. Econ. 19 (4), 25–47.
Jakiela, P., 2011. Social preferences and fairness norms as informal institutions: experimental evidence. Am. Econ. Rev. 101 (3), 509–513.
Keaveny, T.J., Inderrieden, E.J., 2000. Gender differences in pay satisfaction and pay expectations. J. Managerial Issues 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. Negotiat. J. 25 (4), 515–531.
Konow, J., Saijo, T., Akai, K., 2008. Morals and Mores? Experimental Evidence on Equity and Equality from the US and Japan. Technical Report. Institute of Social and Economic Research, Osaka.
Kray, L.J., Thompson, L., 2004. Gender stereotypes and negotiation performance: an examination of theory and research. Res. Organ. Behav. 26, 103–182.
Kugler, K.G., Reif, J.A., Kaschner, T., Brodbeck, F., 2018. Gender differences in the initiation of negotiations: a meta-analysis. Psychol. Bull. 144 (2), 198–222.
Ligon, E., Schechter, L., 2012. Motives for sharing in social networks. J. Dev. Econ. 99 (1), 13–26.
Mazei, J., Hüffmeier, J., Freund, P.A., Stuhlmacher, A.F., Bilke, L., Hertel, G., 2015. A meta-analysis on gender differences in negotiation outcomes and their moderators. Psychol. Bull. 141 (1), 85–104.
McClure, E.B., 2000. A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. Psychol. Bull. 126 (3), 424.
Nash, J.F., 1953. Two-person cooperative games. Econometrica 21 (1), 128–140.
Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Q. J. Econ. 122 (3), 1067–1101.
Ruffle, B.J., Sherman, A., Shtudiner, Z., 2022. Gender and beauty price discrimination in produce markets. J. Behav. Exp. Econ. 97, 101825.
Rutström, E.E., Wilcox, N.T., 2009. Stated beliefs versus inferred beliefs: amethodological inquiry and experimental test. Games Econ. Behav. 67 (2), 616–632.
Shan, W., Keller, J., Joseph, D., 2019. Are men better negotiators everywhere? A meta-analysis of how gender differences in negotiation performance vary across cultures. J. Organ. Behav. 40 (6), 651–675.
Small, D., Gelfand, M., Babcock, L., Gettman, H., 2007. Who goes to the bargaining table? The influence of gender and framing on the initiation of negotiation. J. Pers. Soc. Psychol. 93 (4), 600–613.
Solnick, S., 2001. Gender differences in the ultimatum game. Econ. Inq. 39 (2), 189–200.
Solnick, S.J., Schweitzer, M.E., 1999. The influence of physical attractiveness and gender on ultimatum game decisions. Organ. Behav. Hum. Decis. Process. 79 (3), 199–215.
Verschoor, A., D'Exelle, B., Perez-Viana, B., 2016. Lab and life: does risky choice behaviour observed in experiments reflect that in the real world? J. Econ. Behav. Organ. 128, 134–148.
Vyas, S., Kumaranayake, L., 2006. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan. 21 (6), 459–468.
Wilson, R.K., Eckel, C.C., 2006. Judging a book by its cover: beauty and expectations in the trust game. Polit. Res. Q. 59 (2), 189–202.