Sutter, Matthias; Bosman, Ronald; Kocher, Martin; van Winden, Frans

Working Paper

Gender pairing and bargaining: Beware the same sex!


Provided in Cooperation with:
Institute of Public Finance, University of Innsbruck

Suggested Citation: Sutter, Matthias; Bosman, Ronald; Kocher, Martin; van Winden, Frans (2008) : Gender pairing and bargaining: Beware the same sex!, Working Papers in Economics and Statistics, No. 2008-27

This Version is available at:
http://hdl.handle.net/10419/71915

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Gender pairing and bargaining – Beware the same sex!

Matthias Sutter, Ronald Bosman, Martin Kocher and Frans van Winden

2008-27
Gender pairing and bargaining –
Beware the same sex!*  

Matthias Sutter#, Ronald Bosman§, Martin Kocher†, and
Frans van Winden‡

This version: December 2008

Abstract: We study the influence of gender and gender pairing on economic decision making in an experimental two-person bargaining game where the other party’s gender is known to both actors. We find that (1) gender per se has no significant effect on behavior, whereas (2) gender pairing systematically affects behavior. In particular, we observe much more competition and retaliation and, thus, lower efficiency when the bargaining partners have the same gender than when they have the opposite gender. These findings are consistent with predictions from psychology. Implications of our results for real-world organizations are discussed.

Keywords: gender pairing, bargaining, psychology, experiment.

JEL-classification: C72, C91, C92.

---

* We thank two referees, Jordi Brandts and participants at the Royal Economic Society Meeting 2004 as well as the European Economic Association Meeting 2004 for very useful comments on how to sharpen the focus of the paper. Financial support from the Center of Experimental Economics at the University of Innsbruck (sponsored by Raiffeisen-Landesbank Tirol) and the Austrian Science Foundation (FWF grant P16617) is gratefully acknowledged.

# University of Innsbruck, Department of Public Finance. Universitaetsstr. 15/4, A-6020 Innsbruck, Austria, and University of Gothenburg, Department of Economics. Box 640, SE-40530 Göteborg, Sweden. e-mail: matthias.sutter@uibk.ac.at

§ De Nederlandsche Bank, Financial Stability Division. Westeinde 1, 1017 ZN Amsterdam, The Netherlands. e-mail: r.a.j.bosman@dnb.nl

† Corresponding author. University of Munich, Department of Economics. Geschwister-Scholl-Platz 1, D- 80539 Munich, Germany, and University of Innsbruck, Department of Public Finance. Universitaetsstrasse 15/4, A-6020 Innsbruck, Austria. e-mail: martin.kocher@lrz.uni-muenchen.de

‡ University of Amsterdam, CREED/Department of Economics. Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. e-mail: fvwinden@fee.uva.nl
1 Introduction

Numerous field and laboratory studies have addressed the economic behavior of men and women, finding, for instance, differences in the choice of a profession (Sokoloff, 1992), in salaries and promotions (Ginther and Hayes, 2003), job hiring and firing (Ginther and Kahn, 2004), team behavior (Dufwenberg and Muren, 2006), risk taking (Croson and Gneezy, 2008) or behavior in competitive environments (Gneezy et al., 2003; Gupta et al., 2005; Niederle and Vesterlund, 2007).¹ So, gender² has an impact on the functioning of organizations. Some differences between men and women, in particular those with respect to salaries, promotions and hiring, have been linked to differences in the bargaining behavior of men and women (Watson, 1994; Bowles et al., 2005). However, the evidence on gender differences in bargaining is not fully conclusive, as we will show in greater detail in section 2. This is especially true for many laboratory experiments. The mixed, and frequently insignificant, results on gender and bargaining may be related to different approaches for measuring gender effects. First, field and laboratory studies differ with respect to the degree of control over the structural characteristics of a bargaining situation. Given that structural ambiguity is expected to trigger gender differences in economic behavior (Bowles and McGinn, 2002), possible differences between field and laboratory studies may be traced back to this issue. Second, controlled laboratory experiments differ frequently in whether or not participants know the gender of their bargaining partner. This methodological difference may produce different results. Third, many studies do not control for gender pairing, but only for gender. However, research in psychology suggests that gender pairing has an effect on (economic) behavior.

¹ A comprehensive overview of the literature is provided by Croson and Gneezy (2008).
² We use the expressions „sex“ and „gender“ interchangeably throughout the paper.
It is precisely the latter aspect of the research on gender differences that we are investigating. We present an experimental study where we examine the importance of gender pairing for bargaining behavior. Our vehicle of research is the power-to-take game (see Bosman and van Winden, 2002, Bosman et al., 2005), which is a two-person bargaining game that relates to several important economic situations, such as principle-agent relationships.\footnote{The details of the game and its relevance for economic decision making will be explained in section 3.}

We assess the importance of gender pairing by looking at the four possible combinations of gender in this two-person bargaining game. We find that gender \textit{per se} has no significant effect on behavior, but that gender \textit{pairing} has a strong influence. In particular, we observe much more competition and retaliation and, thus, lower efficiency when the bargaining partners have the same gender than when they have the opposite gender.

The rest of the paper is organized as follows: In section 2 we will give a brief account of previous economic studies on the effects of gender, and in particular of gender pairing, in bargaining. Section 3 describes the power-to-take game and the motivation for using this game. It also presents our hypotheses on the effects of gender and gender pairing based on findings in social and evolutionary psychology. Section 4 is devoted to the experimental design, while the results are given in section 5. Section 6 concludes with a discussion of the implications of our findings with respect to applied organizational research.

\section{Gender and bargaining}

Numerous field studies have addressed the possible effects of gender in the context of bargaining. Ayres (1991) and Ayres and Siegelman (1995) are two prominent examples of a controlled field experiment. They examine the bargaining behavior of men and women in negotiations for the purchase of a new car. They find that women get worse deals from women than from men, which implies that gender pairing is important for bargaining. The
meta-analyses of the influence of gender on bargaining outcomes by Walters et al. (1998) and Stuhlmacher and Walters (1999) suggest that men earn more in negotiations than women, even though the difference is rather small in economic terms. Craver and Barnes (1999), however, claim that there are no statistically significant differences in negotiation outcomes and performances between men and women. The problem with field studies is the fact that they are highly context-dependent. Robertson (2001), for instance, shows that gender differences in salaries depend upon the degree of regulation in an industry and the transparency of appropriate salary standards. Hence, the field evidence for the claim that women are worse bargainers than men is non-conclusive. This raises the question whether controlled laboratory experiments provide less ambiguous evidence.

The experimental dictator game provides a good starting point. Since the dictator game is basically an individual decision making task where an individual has to allocate a sum of money between him- or herself and one other person, it eliminates possibly confounding factors of strategic interaction like risk aversion which might affect men and women differently. To date, the evidence on gender effects in the dictator game is ambiguous, though. Whereas Bolton and Katok (1995), Frey and Bohnet (1995) and Carpenter et al. (2005) find no evidence for gender differences, Eckel and Grossman (1998) and Fehr et al. (2006) report women to be significantly less selfish than men. Concerning the influence of gender pairing, Ben-Ner et al. (2004) find that women give significantly less to women than to men and persons of unknown gender.

In order to study bargaining behavior in a real interactive environment, the ultimatum game is a more suitable tool. In this game, the proposer can offer an amount $x \leq E$ to a responder. If the responder accepts, the proposer earns $E - x$, and the responder $x$. If the responder rejects, both earn nothing. With respect to gender pairing, two studies seem

---

4 Camerer (2003) provides a brief overview of experimental studies on the influence of gender on bargaining behavior.
relevant. Eckel and Grossman (2001) show that women are more cooperative than men in a repeated ultimatum game where proposers and responders face each other. Whereas gender seems to play a role *per se* in determining bargaining behavior, Eckel and Grossman note that gender pairing is also important. In particular, women paired with women almost never fail to reach an agreement, which they interpret as solidarity. Solnick (2001), however, finds the opposite effects in a one-shot ultimatum game using the strategy method: Women making offers to women face the highest rejection rates.\(^5\)\(^6\) One explanation for the different findings might be differences in the experimental procedure, though. In Solnick’s (2001) study participants sat in cubicles when making their decision and had no visual contact with their bargaining partners. This is in contrast to the experiment of Eckel and Grossman (2001) where proposers and responders sat opposite each other and faced each other.\(^7\) With such a design, the effects of gender and gender pairing might easily be confounded with the effects of visual expression or beauty.\(^8\)

To summarize, the evidence on the role of gender and, in particular, of gender pairing in bargaining, both from field studies and experimental studies, is non-conclusive. It is not easy

---

\(^5\) Holm (2000) reports a general tendency of both sexes to discriminate against women in a coordination game (the battle of the sexes game), which is, however, not directly comparable to bargaining games.

\(^6\) The experimental evidence in another bargaining game, the trust game, is also mixed. Croson and Buchan (1999) find in their cross-cultural study that women show more reciprocal behavior than men. However, Fehr et al. (2003) do not find any gender difference using a representative sample of the German population. Sutter and Kocher (2007) report also no gender differences in their trust game study, where they had participants from various age groups, ranging from 8-year old children to 80-year old retired persons. The effects of gender pairing could not be assessed in their study because the gender of interacting partners was not revealed.

\(^7\) More precisely, four proposers sat opposite four responders. Participants were told that they would be paired with one of the opposite (four) bargaining partners.

\(^8\) Schweitzer and Solnick (1999), for instance, show in an ultimatum game that there is something like a beauty premium, meaning that more attractive people are offered more. Frey and Bohnet (1999a, 1999b) find that the mere identification of bargaining ‘partners’ leads to more cooperative bargaining behavior. Hence, it might have been identification rather than gender that drives the results of Eckel and Grossman (2001) into another direction than the results of Solnick (2001). Eckel and Grossman (2008) discuss other possible sources for the different results between Eckel and Grossman (2001) and Solnick (2001) at greater length.
to explain why the evidence is so mixed. Studies differ in many important methodological ways, namely the way gender pairing is controlled for, the way in which the bargaining game is implemented, or the way in which (or whether) gender is revealed. The studies mentioned above focus mainly on gender effects. In our view, however, gender pairing effects are equally important since in real life individuals typically know the gender of people with whom they are interacting. In the following we, therefore, present a controlled experiment on gender pairing effects.

3 The power-to-take game and hypotheses from psychology

3.1 The power-to-take game

The power-to-take game is a two-person, two-stage game between a ‘take authority’ (with endowment $E_{take}$) and a ‘responder’ (with $E_{resp}$). In the first stage, the take authority decides on a so-called take rate $t \in [0,1]$, which determines the part of the responder’s endowment after the second stage that will be transferred to the take authority. In the second stage, the responder can decide to destroy a part $d$ of $E_{resp}$, with $d \in [0,1]$. For the take authority the payoff is thus given by $E_{take} + t(1-d)E_{resp}$. For the responder, the payoff equals $(1-t)(1-d)E_{resp}$.

Even though the power-to-take game is very simple, its structure resembles a broad range of economic situations. First of all, by its very nature it is a bargaining game with two parties having influence on the economic surplus (of the responder) which can be distributed between both parties. The game can be interpreted as a principal-agent relationship. The principal can be seen as the take authority who decides on the incentive scheme for the agent (the responder). The scheme involves a claim on the value product that can be generated by the working capital that the agent has at his disposal. If offended by the scheme, the agent
may feel urged to punish the principal by producing less value, which is also costly for the agent when it conflicts with the material incentives provided by the scheme. Another example of the economic relevance of the power-to-take game is monopolistic pricing. The price selected by the firm entails a claim on the consumer surplus. If the buyer feels that the price is outrageous, buyers may be induced to punish the firm by buying less than the rational ‘text book’-buyer would do.

Compared to simpler games – like the ultimatum game – the power-to-take game has a richer structure. The ultimatum game with its all-or-nothing decision of the responder is less general than the power-to-take game. Hence, the possibility of (almost) continuous destruction rates allows for more variability concerning the efficiency of an interaction. The all-or-nothing feature of the ultimatum game is in particular responsible for the modal offer almost always being a 50:50 split between proposer and responder. The fine-tuning of destruction rates has produced a much larger variability of takes rates in the power-to-take game (Bosman and van Winden, 2002; Bosman et al., 2005). Compared to the ultimatum game, the power-to-take game also has a rather asymmetric distribution of power since the take authority’s endowment is not at stake. Hence, there is a much more distinctive power- relation in the power-to-take game. The asymmetry following from that seems a realistic feature in many real-life bargaining processes.

3.2 Hypotheses

Obviously, traditional economic theory with its assumption of rational, money- maximizing agents is unable to explain gender differences, because it predicts both men and women to choose the same optimizing behavior when they are in the same situation. Models of economic behavior in the presence of social preferences (e.g., Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002) cannot predict any differences
between men and women either, given that men and women are assumed not to have systematically different social preferences. Hence, we resort to psychology in order to put forward hypotheses on gender-related behavior in the power-to-take game.

3.2.1 Social psychology

Research in social psychology suggests that gender differences are a situational phenomenon rather than a stable and universal trait of the behavior of men and women. Social role theory (Snyder and Ickes, 1985) and status characteristics theory (Carli and Eagly, 1999; Eagly et al., 2000) expect gender differences to emerge either when men and women differ with respect to their social position, in particular their social status\(^9\) and social role, or when the situation is not highly structured. The latter aspect is also referred to as the structural ambiguity of a situation (Bowles and McGinn, 2002), meaning the degree of potential variation in a party’s interpretation of the economic structure of the bargaining situation. When the situation is unambiguous, then the prediction would be that there are no gender differences. Since the power-to-take game is a highly structured game, our first hypothesis is that we expect no gender differences with respect to take rates and destruction rates.

3.2.2 Evolutionary psychology

Evolutionary psychology explains human behavior as an adaptation to two primary challenges of humans: survival and reproductive success (Buss, 1999). Even though males and females have adapted differently to these challenges, reproductive success has influenced behavior

---

\(^9\) Ball et al. (2001) conduct an experiment on the economic consequences of social status. In a market game, participants with higher status (either randomly assigned or “earned” through a quiz) got more favourable prices as a seller or a buyer.
towards members of the own sex and the opposite sex in a systematic way. Trivers’ (1972) theory of parental investment and sexual selection predicts that, as a consequence of the competition for a mate, rivalry and aggression in behavior should be more intense within the same sex (intra-sexual competition) than against the opposite sex (intersexual competition).\textsuperscript{10} This is quite natural given that the members of one’s own sex are the primary competitors for valuable members of the opposite sex. Applied to the power-to-take game, evolutionary psychology seems to predict that interaction between members of the same sex will be more aggressive or competitive.\textsuperscript{11} Therefore, we put forward as our second hypothesis that take rates and destruction rates are to be higher under same gender pairing than under mixed gender pairing.

4 Experimental design

At the beginning of the experiment, participants received a show up fee of 60 ECU (experimental currency units) (worth 4.5€) and an initial endowment of $E_{\text{take}} = E_{\text{resp}} = 120$.

\textsuperscript{10} It is noteworthy that the empirical evidence on acts of aggression supports the predictions from evolutionary psychology. Verbal as well as physical aggression is more frequently directed against members of the same gender than against members of the opposite gender (Hyde, 1984, 1986; Campbell, 1995). The other side of the coin is that behavior towards members of the opposite gender is typically more cooperative. At first sight, evolutionary psychology might also have implications for gender differences \textit{per se}, such that males are typically the more dominant or aggressive sex (Trivers, 1972; Archer, 1996). Greater male aggressiveness has especially been evidenced by data on physical aggression and criminal offenses like homicides (Knight et al., 1996). However, the evidence on low-key (verbal) aggression is far less conclusive (Kinney, 2001; Ramirez et al., 2001). In particular, experimental studies – in comparison to field studies – tend to find significantly less gender differences in aggression. According to Fischer and Rodriguez Mosquera (2001, p. 19), “experimental research has suggested that men and women are equally aggressive, if concerns and appraisals are rendered equal for men and women”. Since concerns and appraisals for money – the key motivation to participate in experiments for both men and women – are to be expected equal for men and women, it is difficult to derive from evolutionary psychology any hypothesis on the influence of gender \textit{per se} on bargaining.

\textsuperscript{11} Recently, several papers have investigated the effects of gender on competitive behavior. See, e.g., Gneezy et al. (2003), Niederle and Vesterlund (2007), or Gneezy et al. (2008).
ECU (worth 9€). Take rates $t$ and destruction rates $d$ could be chosen in integer percentages. Assuming maximization of own payoffs, a take rate of $t = 99\%$ and a destruction rate of $d = 0\%$ would be a subgame-perfect Nash equilibrium outcome.\textsuperscript{12} Note that only if $t = d = 0\%$, experimental earnings of both players would be equal. In all other cases, the responder always earns less than the take authority.\textsuperscript{13}

In order to assess the influence of expectations, we requested responders to indicate the expected take rate before they got to know the actual one. Likewise, we asked take authorities for the expected destruction rate after having decided on the take rate and before being informed about the actual destruction rate.\textsuperscript{14}

Our four different treatments (FF, FM, MF, MM) result from a $2 \times 2$ matrix determined by the take authority’s and the responder’s gender in a between-subject design. Subjects were informed about the gender of both roles in the instructions in the following way (see appendix A2):\textsuperscript{15} When introducing the roles A (take authority) and B (responder), we inserted a single sentence stating the gender of the subject in each role. For example, in the female-male treatment (FM), this sentence ran as follows: “The subject in the role of A is a woman, and

\textsuperscript{12} t = 100\%$ and $d = 0\%$ constitute also a Nash equilibrium. However, in this case $d = 0\%$ is only a weakly dominant strategy for the responder, since every other feasible choice of $d$ yields the same final payoff of zero for the responder. Only if $t < 100\%, d = 0\%$ is a strictly dominant strategy for the responder.

\textsuperscript{13} Recall that the responder can only destroy his or her own income ($E_{\text{resp}}$), but not that of the take authority ($E_{\text{take}}$).

\textsuperscript{14} We did not pay for the accuracy of expectations. Readers may be concerned about the lack of financial incentives for reporting expectations. There is, however, evidence that providing financial incentives for probability estimates does not change the data much: “When one examines subjects’ choices and decisions the observed effects of financial incentives were with one exception not dramatic. Subjects with financial incentives appeared to perform somewhat better than their counterparts without such incentives, but the differences were not great, were generally not statistically significant and did not hold in every case” (Grether, 1992, p. 54; see also Camerer and Hogarth, 1999).

\textsuperscript{15} The game was framed as neutral as possible, avoiding any suggestive terms like take authority or take rate.
the subject in the role of B is a man.” 16 Nowhere else did we emphasize the role of gender in the game.

The experiment was computerized with the help of z-Tree (Fischbacher, 2007). For each treatment we got 19 pairs. 17 About 75% of our 152 participants were undergraduate students of economics or business administration. Most of the rest was enrolled in medicine or psychology. Sessions lasted less than 50 minutes, with participants earning in total an average of 162 ECU (about € 12).

5 Results

We divide our data analysis into two parts (see Table A1 in the appendix for raw data). First, we analyze behavior and test our hypotheses. Thereafter, we explore the relation between behavior and expectations.

5.1 Take rates and destruction rates

Table 1 shows averages and standard deviations of take rates and destruction rates for each of the four treatments (with \( N = 19 \) in each treatment). Frequencies of destruction are calculated by classifying responder behavior with \( d > 0 \) as destruction.

Table 1 about here

---

16 We could also have stated the first name of the respective bargaining partner. But note that Holm (2000) has shown in a coordination game that experimental results were not significantly different under the following two conditions: (a) Subjects knew the gender of the bargaining partner. (b) Subjects knew the first name of the bargaining partner. Hence, we decided against using first names to avoid potential violation of anonymity.

17 All sessions were conducted by the same experimenter.
Averaging over all treatments (see the outer right column of Table 1), the take rate equals two thirds of the responder’s endowment $E_{\text{resp}}$. Responders destroy on average 30% of their initial endowment, with about 45% of the responders destroying at least some amount of money (i.e. $d > 0$). Due to the fact that the take authorities’ endowment $E_{\text{take}}$ is not at stake, take authorities earn on average considerably more than responders (230 ECU vs. 94 ECU). Looking at single treatments, take rates are highest in the FF-treatment (75%), where females face females. Average destruction rates (46%) and the frequency of destruction (63%) are highest in the MM-treatment, where males interact with males.

In order to test our first hypothesis on the effects of gender per se, we aggregate treatments by the gender of the decision maker. For example, female take rates are derived from treatments FF and FM, while female destruction rates consider treatments FF and MF. The left-hand side of Table 2 reports the relevant figures. Take rates are 69.66% for female take authorities and 63.68% for male take authorities. Destruction rates are on average at 30% for females and males. In sum, we do not find any significant differences between females and males with respect to take rates, destruction rates or frequencies of destruction (with all $p$-values larger than 0.2). This holds true when we compare male and female behavior across all treatments, but also when we compare on a more disaggregated level treatments FF and MM, respectively FM and MF.\(^{18}\) Hence, we cannot reject our first hypothesis that predicts no gender differences per se.

\(^{18}\) The absence of gender effects is based on the fact that both men and women are tough to their own gender and softer to the other gender. However, the absence of gender effects for the destruction rates comes about in a slightly different way. Men are very tough to men and much softer to women while women are tougher to women than to men, but the difference is considerable smaller than the one for men.
In order to test our second hypothesis on the influence of gender pairing, we control for
gender when comparing decisions in treatments with same gender pairing, respectively mixed
gender pairing (please refer to Table 1). Given that the take authority is female, we find
significantly higher take rates when the responder is female (FF: 75%) than when the
responder is male (FM: 64%) \( (p < 0.05; \) one-sided Mann-Whitney U-test\(^{19}\). The same effects
of gender pairing can be found for male take authorities, with higher take rates in MM (70%)
than in MF (57%) \( (p = 0.06; \) one-sided Mann-Whitney U-test\(^{20}\).

Holding the responder’s gender constant, we find that the destruction rate is
significantly larger if a male responder is paired with a male take authority (MM: 46%) rather
than a female take authority (FM: 13%) \( (p < 0.01; \) one-sided U-test). The frequency of
destruction is also significantly larger in MM than in FM \( (p < 0.01; \) one-sided \( \chi^2 \)-test). For
female responders, gender pairing has no significant effect on destruction rates and the
frequency of destruction.

Another way to show the effects of gender pairing is to pool treatments by gender
pairing, as is done on the right-hand side of Table 2. Treatments FF and MM are pooled to
‘same gender pairing’, and FM and MF to ‘mixed gender pairing’\(^{21}\). Take rates, destruction
rates and the frequency of destruction are always significantly higher under same gender
pairing than under mixed gender pairing, as can be discerned from the significance levels on
the right-hand side of Table 2. Take rates are about 20% higher when subjects face the same
gender than when they face the opposite gender. Destruction rates with same gender pairing
are more than double the corresponding values for mixed gender pairing, and the frequency of

\(^{19}\) Since our second hypothesis provides a directional prediction concerning the effects of gender pairing on take
rates or destruction rates, we can apply one-sided tests in the statistical analysis.

\(^{20}\) We also find significantly higher take rates in FF than in MF \( (p = 0.01; \) one-sided Mann-Whitney U-test). All
other pairwise comparisons yield no significant differences.

\(^{21}\) Pooling is possible, because take rates, destruction rates and frequencies of destruction do not differ in
medians (Mann-Whitney U-test) nor in the distribution of values (Kolmogorov-Smirnov-test) between FF and
MM (same gender pairing), nor between FM and MF (mixed gender pairing).
destruction is about 80% larger. Remarkably, under same gender pairing, 10 out of 38 decision makers chose \( t > 95\% \), whereas this occurs only twice under mixed gender pairing \((p < 0.02; \chi^2 = 6.33; \text{one-sided})\). Regarding the destruction rates, 10 decision makers in the same gender pairing condition chose \( d > 95\% \), but only 4 decision makers in the mixed gender pairing condition \((p < 0.05; \chi^2 = 3.15; \text{one-sided})\).

Table 3 sheds light on the influence of gender pairing from another perspective. It reports average destruction rates for different intervals of the take rate. With the exception of the interval [81\%, 90\%], average destruction rates are always higher under same gender pairing than under mixed gender pairing.\(^{22}\) In sum, we have found strong support for our second hypothesis, which states that take rates and destruction rates are higher under same gender pairing than under mixed gender pairing.

5.2 Expectations

5.2.1 Expected versus actual decisions

Table 4 reports expected take and destruction rates and compares them to actual decisions. Expected take rates are significantly smaller than the actual ones in each single treatment, falling, on average, 22 percentage points short of the actual take rate \((p < 0.01 \text{ in FF}, p < 0.05 \text{ in FM}, p < 0.1 \text{ in MF and MM}; \text{two-sided Wilcoxon signed ranks-test})\). Interestingly, expected take rates do not differ significantly between any two treatments, nor do they depend on gender or gender pairing. This may have been a consequence of

\(^{22}\) In the interval [0\%, 10\%], we cannot compare the two conditions because for the same gender pairing we have no observation in this interval.
expectations not having been incentivized, even though Grether (1992) indicates that non-incentivicing need not have an effect. It could also be that responders did not anticipate that take authorities of the same gender behave more aggressively.\textsuperscript{23}

\textit{Table 4 about here}

Destruction rates expected by female take authorities (in treatments FF and FM) are not significantly different from actual destruction rates, suggesting that female take authorities have a good intuition of which destruction rates will be evoked by their specific take rates. However, male take authorities (in MF and MM) expect significantly lower destruction rates than their counterpart responders actually choose ($p < 0.05$; two-sided Wilcoxon signed ranks test). Comparing expected destruction rates across treatments we find no significant difference in any pairwise comparison.

5.2.2 The influence of expected take rates on the likelihood of destruction

Figure 1 plots individual data on the take rates expected by responders (on the horizontal axis) versus the actual take rates chosen by take authorities (on the vertical axis). Points above (below) the diagonal indicate that expectations were lower (higher) than actual decisions, and, thus, too optimistic (pessimistic). We have marked those responders who destroyed parts or all of their endowment by a cross. The frequency of points lying above or below the diagonal is significantly different between responders who destroyed something or

\textsuperscript{23} This result might be explained by the well-known hot-cold empathy gap (Loewenstein, 2000) which states that people are bad in predicting behavior in a state they are not currently in themselves. Hence, men (respectively women) might not be able to put themselves ‘into the shoes’ of a male and female take authority, therefore expecting the same behavior of both male and female take authorities.
everything and those who destroyed nothing ($p < 0.05$; $\chi^2 = 3.15$; one-sided test). This suggests that those disappointed by the take rate are more likely to destroy.

6 Conclusion

Our results on behavior in a power-to-take game experiment suggest that gender pairing is an important determinant in bilateral relationships. In particular, we have found that take authorities demand significantly more from responders of the same gender. In turn, responders’ destruction rates are higher when they deal with a take authority of their own gender. No differences were found between inter-male and inter-female bargaining. Furthermore, we find no significant differences in behavior between men and women per se. Overall, our results are in line with predictions derived from social and evolutionary psychology.

When comparing our findings with previous experimental studies, we would like to stress that the existing evidence on the influence of gender per se in two-person bargaining games, like the ultimatum game or the dictator game, is not fully conclusive (see Camerer, 2003, and Croson and Gneezy, 2008, for surveys). Even though there are some studies indicating that men perform better in bargaining and that women are more cooperative (and thus easier to exploit), there is also counter-evidence. Besides, it is likely that many papers on two-person bargaining do not report the effects of gender on bargaining, because they find no statistically significant difference. The inclination to report (and publish) only significant results may lead to a greater emphasis on gender differences than is actually justified.

Perhaps more importantly, only few studies have controlled for gender pairing when studying the effects of gender per se. Gender differences found in the literature may actually
vanish if results were controlled for gender pairing. Note, for instance, that if we had run only treatments FF and MF in our experiment, we could have reported significant differences in take rates between women (75%) and men (57%). Controlling for gender pairing, we have found no differences at all (neither in the same gender pairing condition, where we compared FF with MM, nor in the mixed gender pairing condition, comparing FM with MF).

Our results have implications for bargaining processes or principal-agent relationships in organizations, since men and women apparently behave differently depending upon whom they are interacting with. As a consequence, it may be in the interest of an institution (like an organizational unit within a firm) involved in bargaining to strategically select the gender of its representative. Our results indicate that mixed gender pairing fosters more cooperation and entails a lower probability of an inefficient outcome. Same gender pairing leads to more competitive behavior but also to a higher likelihood that scarce resources will be wasted.

\[24\text{ A methodological implication of our results is that (both field and experimental) studies of behavioral differences between men and women should control for gender pairing and that failing to do so might lead to seriously misleading conclusions.}\]
References


Tables and Figures

**Table 1. Decisions**

<table>
<thead>
<tr>
<th>treatment*</th>
<th>FF</th>
<th>FM</th>
<th>MF</th>
<th>MM</th>
<th>overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>take rate (%)</td>
<td>75.42*</td>
<td>63.89*</td>
<td>57.16*</td>
<td>70.21*</td>
<td>66.67</td>
</tr>
<tr>
<td>(standard deviation)</td>
<td>(20.13)</td>
<td>(15.72)</td>
<td>(25.24)</td>
<td>(25.56)</td>
<td>(22.65)</td>
</tr>
<tr>
<td>destruction rate (%)</td>
<td>36.63*</td>
<td>13.42*</td>
<td>24.32*</td>
<td>45.84*</td>
<td>30.05</td>
</tr>
<tr>
<td>(standard deviation)</td>
<td>(43.62)</td>
<td>(31.71)</td>
<td>(35.29)</td>
<td>(41.94)</td>
<td>(39.64)</td>
</tr>
<tr>
<td>frequency of destruction (%)</td>
<td>52.63*</td>
<td>21.05*</td>
<td>42.11</td>
<td>63.16*</td>
<td>44.74</td>
</tr>
<tr>
<td>profit take authority*</td>
<td>ECU 230.54</td>
<td>243.76</td>
<td>228.72</td>
<td>218.36</td>
<td>230.34</td>
</tr>
<tr>
<td>profit responder*</td>
<td>ECU 85.50</td>
<td>100.31</td>
<td>102.10</td>
<td>86.63</td>
<td>93.64</td>
</tr>
</tbody>
</table>

* FF: both roles females; FM (MF): female (male) take authorities, male (female) responders; MM: both roles males.

1,2,3 significant at $p < 0.05$ (one-sided-test)

4 significant at $p < 0.07$ (one-sided-test)

$N = 19$ for each single treatment.

**Table 2. Decisions grouped by gender and gender pairing**

<table>
<thead>
<tr>
<th>gender pairing</th>
<th>females</th>
<th>males</th>
<th>significance</th>
<th>same</th>
<th>mixed</th>
<th>significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>take rate (%)</td>
<td>69.66</td>
<td>63.68</td>
<td>n.s.</td>
<td>72.82</td>
<td>60.53</td>
<td>$p &lt; 0.01$ (one-sided U-test)</td>
</tr>
<tr>
<td>destruction rate (%)</td>
<td>30.47</td>
<td>29.63</td>
<td>n.s.</td>
<td>41.24</td>
<td>18.87</td>
<td>$p &lt; 0.01$ (one-sided U-test)</td>
</tr>
<tr>
<td>frequency of destruction (%)</td>
<td>47.37</td>
<td>42.11</td>
<td>n.s.</td>
<td>57.89</td>
<td>31.58</td>
<td>$p &lt; 0.02$ (one-sided $\chi^2$-test)</td>
</tr>
</tbody>
</table>

n.s. not significant.
Table 3. Take rates and destruction rates

<table>
<thead>
<tr>
<th>take rate</th>
<th>same gender destruction rate (average)</th>
<th>N</th>
<th>mixed gender destruction rate (average)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-10%</td>
<td>-</td>
<td>0</td>
<td>50.0</td>
<td>1</td>
</tr>
<tr>
<td>11-20%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>21-30%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>31-40%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>41-50%</td>
<td>15.7</td>
<td>7</td>
<td>6.5</td>
<td>13</td>
</tr>
<tr>
<td>51-60%</td>
<td>50.0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>61-70%</td>
<td>28.1</td>
<td>9</td>
<td>15.3</td>
<td>6</td>
</tr>
<tr>
<td>71-80%</td>
<td>48.6</td>
<td>5</td>
<td>31.3</td>
<td>8</td>
</tr>
<tr>
<td>81-90%</td>
<td>33.0</td>
<td>3</td>
<td>46.7</td>
<td>3</td>
</tr>
<tr>
<td>91-100%</td>
<td>83.0</td>
<td>10</td>
<td>50.0</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4. Expected take rates and destruction rates versus actual decisions

<table>
<thead>
<tr>
<th>treatment</th>
<th>FF</th>
<th>FM</th>
<th>MF</th>
<th>MM</th>
<th>overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>expected take rate in % (average)</td>
<td>41.58</td>
<td>44.42</td>
<td>39.21</td>
<td>50.53</td>
<td>43.93</td>
</tr>
<tr>
<td>(standard deviation)</td>
<td>(33.08)</td>
<td>(29.42)</td>
<td>(26.84)</td>
<td>(24.26)</td>
<td>(28.33)</td>
</tr>
<tr>
<td>actual take rate in %</td>
<td>75.42</td>
<td>63.89</td>
<td>57.16</td>
<td>70.21</td>
<td>66.67</td>
</tr>
<tr>
<td>expected destruction rate in % (average)</td>
<td>23.95</td>
<td>16.58</td>
<td>5.26</td>
<td>19.58</td>
<td>16.34</td>
</tr>
<tr>
<td>(standard deviation)</td>
<td>(33.69)</td>
<td>(29.06)</td>
<td>(9.79)</td>
<td>(32.37)</td>
<td>(28.25)</td>
</tr>
<tr>
<td>actual destruction rate in %</td>
<td>36.63</td>
<td>13.42</td>
<td>24.32</td>
<td>45.84</td>
<td>30.05</td>
</tr>
<tr>
<td>expected frequency of destruction in %</td>
<td>52.63</td>
<td>36.84</td>
<td>31.58</td>
<td>47.37</td>
<td>42.11</td>
</tr>
<tr>
<td>actual frequency of destruction in %</td>
<td>52.63</td>
<td>21.05</td>
<td>42.11</td>
<td>63.16</td>
<td>44.74</td>
</tr>
</tbody>
</table>
Figure 1. Actual vs. expected take rate and destruction (N = 76)
Appendix

A1: Individual raw data

Table A1. Individual data

<table>
<thead>
<tr>
<th>female</th>
<th>female vs. female</th>
<th>destruction rate</th>
<th>expected take rate</th>
<th>expected destruction rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>pair</td>
<td>take rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>45</td>
<td>3</td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>0</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>0</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0</td>
<td>50</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>70</td>
<td>30</td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>70</td>
<td>0</td>
<td>80</td>
<td>40</td>
</tr>
<tr>
<td>10</td>
<td>80</td>
<td>65</td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>80</td>
<td>0</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>12</td>
<td>80</td>
<td>98</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>88</td>
<td>0</td>
<td>10</td>
<td>60</td>
</tr>
<tr>
<td>14</td>
<td>90</td>
<td>0</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>16</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>17</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>50</td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>100</td>
<td>100</td>
<td>60</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>male vs. female</th>
<th>male vs. male</th>
<th>destruction rate</th>
<th>expected take rate</th>
<th>expected destruction rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>pair</td>
<td>take rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>50</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>0</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>0</td>
<td>75</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>45</td>
<td>60</td>
<td>60</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>0</td>
<td>100</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>0</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>0</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>0</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>65</td>
<td>65</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>70</td>
<td>27</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>75</td>
<td>50</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>75</td>
<td>0</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>75</td>
<td>100</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>90</td>
<td>10</td>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td>18</td>
<td>99</td>
<td>0</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>100</td>
<td>100</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>
A2: Instructions for the FF-treatment (originally in German)

Instructions

Show up fee
Each participant in this experiment receives a show up fee of 60 Experimental Currency Units (ECU), worth 4.5€. You will receive the show up fee irrespective of your decisions in the experiment.

Initial endowment
Each participant receives an endowment of 120 ECU (worth 9€).

Two phases
The experiment consists of two phases. In phase 1 only participant A must make a decision whereas in phase 2 only participant B must make a decision. Every participant thus makes one decision. There is no repetition.
Both, participant A as well as participant B, are female in this experiment. The pairing is random. You will not be informed about whom you were paired with. Your decisions remain anonymous.

Phase 1: participant A chooses a percentage
Participant A must choose a percentage. This percentage determines how much of participant B’s endowment after phase 2 will be transferred to participant A. The percentage chosen by participant A must be an integer in the interval [0, 100]. Zero and one hundred are also possible.

Phase 2: participant B chooses a percentage
Participant B must choose a percentage. This percentage determines which part of participant B’s initial endowment shall be destroyed. The percentage chosen by participant B must be an integer in the interval [0, 100]. Zero and one hundred are also possible.
The transfer from participant B to participant A will be based on the endowment of participant B that is left. Note that the transfer equals the percentage chosen by participant A of the endowment of participant B that is left after phase 2.

Example how to determine one’s payoffs
We will now give an example for the purpose of illustration. As you know both participant A and participant B have an initial endowment of 120 ECU.
Suppose participant A decides that 60% of the endowment of participant B will be transferred to her (participant A). Suppose participant B decides to destroy zero percent of her endowment.
The transfer from B to A is then equal to 72 ECU (60% of 120 ECU).
The total payoff for A at the end of the experiment is equal to 252 ECU (namely, the show-up fee of 60 ECU plus the initial endowment of 120 ECU plus the transfer of 72 ECU).
The total payoff for B at the end of the experiment is equal to 108 ECU (namely, the show-up fee of 60 ECU plus the endowment of 60 ECU minus the transfer of 36 ECU).

Now suppose that in this example participant B had decided to destroy 50% of her own endowment. In this case the transfer from B to A would be 36 ECU (namely, 60% of the remaining endowment of participant B after phase II, which is 60% of 60 ECU).
The total payoff for A at the end of the experiment is equal to 216 ECU (namely, the show-up fee of 60 ECU plus the endowment of 120 ECU plus the transfer of 36 ECU) and for participant B 84 ECU (namely, the show-up fee of 60 ECU plus the remaining endowment of 60 ECU after destruction minus the transfer of 36 ECU).

Summary
In phase 1, each participant A will be randomly paired with a participant B. Pairing remains anonymous throughout the experiment as well as after the experiment.
In phase 1, participant A decides on a percentage that indicates how much of participant B’s endowment will be transferred to participant A. In phase 2, participant B decides which percentage of the initial endowment will be destroyed. From the remaining endowment, participant B has to transfer the percentage chosen by participant A to participant A.

**Other information**

*Calculator*

For your convenience, we have put a calculator on your desk. You can use it in case you want to calculate something.

*The payment procedure*

You will receive your earnings immediately after the end of the experiment from a cashier who is not present during the experiment.

*Exercises*

We ask you to do two exercises in order to become familiar with the procedure. These exercises consist of determining payoffs for fictitious situations. You are not actually paired with another participant during these exercises. Your earnings in these exercises will not be paid out to you. When the exercises have been finished, you have the opportunity to ask questions again. After this the experiment will start.

*Finally*

Please remain silent throughout the experiment. At the end, you are asked to proceed to the cashier one by one and leave the laboratory after receiving your payment.

---

We did not refer to the elicitation of expectations in the experimental instructions. In the computerized program (using z-Tree by Fischbacher, 2007) we asked take authorities after their decision on the take rate to answer the following question: “Which percentage do you expect participant B to destroy?” Responders were asked the following question before learning the actual take rate: “Which percentage do you expect participant A to choose?”
2008-27 Matthias Sutter, Ronald Bosman, Martin Kocher and Frans van Winden: Gender pairing and bargaining – Beware the same sex!

2008-26 Jesus Crespo Cuaresma, Gernot Doppelhofer and Martin Feldkircher: The Determinants of Economic Growth in European Regions.

2008-25 Maria Fernanda Rivas and Matthias Sutter: The dos and don’ts of leadership in sequential public goods experiments.

2008-24 Jesus Crespo Cuaresma, Harald Oberhofer and Paul Raschky: Oil and the duration of dictatorships.


2008-22 Francesco Feri, Bernd Irlenbusch and Matthias Sutter: Efficiency Gains from Team-Based Coordination – Large-Scale Experimental Evidence.


2008-20 Matthias Sutter, Jürgen Huber and Michael Kirchler: Bubbles and information: An experiment.


2008-18 Jürgen Huber and Michael Kirchler: Corporate Campaign Contributions as a Predictor for Abnormal Stock Returns after Presidential Elections.

2008-17 Wolfgang Brunauer, Stefan Lang, Peter Wechselberger and Sven Bienert: Additive Hedonic Regression Models with Spatial Scaling Factors: An Application for Rents in Vienna.

2008-16 Harald Oberhofer, Tassilo Philippovich: Distance Matters! Evidence from Professional Team Sports.

2008-15 Maria Fernanda Rivas and Matthias Sutter: Wage dispersion and workers’ effort.


2008-13 Sergio Currarini and Francesco Feri: Information Sharing Networks in Oligopoly.

2008-12 Andrea M. Leiter: Age effects in monetary valuation of mortality risks - The relevance of individual risk exposure.


2008-10 Harald Oberhofer and Michael Pfaffermayr: Firm Growth in Multinational Corporate Groups.

2008-09 Michael Pfaffermayr, Matthias Stöckl and Hannes Winner: Capital Structure, Corporate Taxation and Firm Age.

2008-08 Jesus Crespo Cuaresma and Andreas Breitenfellner: Crude Oil Prices and the Euro-Dollar Exchange Rate: A Forecasting Exercise.

2008-07 Matthias Sutter, Stefan Haigner and Martin Kocher: Choosing the carrot or the stick? – Endogenous institutional choice in social dilemma situations.


2008-03 Jesus Crespo Cuaresma and Tomas Slacik: On the determinants of currency crises: The role of model uncertainty.

2008-02 Francesco Feri: Information, Social Mobility and the Demand for Redistribution.


2007-31 Andreas Exenberger and Simon Hartmann: The Dark Side of Globalization. The Vicious Cycle of Exploitation from World Market Integration: Lesson from the Congo.


2007-24 Jesus Crespo Cuaresma: Forecasting euro exchange rates: How much does model averaging help?


2007-15 Sergio Currárini and Francesco Feri: Bilateral information sharing in oligopoly.

2007-14 Francesco Feri: Network formation with endogenous decay.

2007-12  **Wolfgang Luhan, Martin Kocher and Matthias Sutter:** Group polarization in the team dictator game reconsidered. *Revised version forthcoming in: Experimental Economics.*

2007-11  **Onno Hoffmeister and Reimund Schwarze:** The winding road to industrial safety. Evidence on the effects of environmental liability on accident prevention in Germany.

2007-10  **Jesus Crespo Cuaresma and Tomas Slacik:** An “almost-too-late” warning mechanism for currency crises.

2007-09  **Jesus Crespo Cuaresma, Neil Foster and Johann Scharler:** Barriers to technology adoption, international R&D spillovers and growth.

2007-08  **Andreas Brezger and Stefan Lang:** Simultaneous probability statements for Bayesian P-splines.

2007-07  **Georg Meran and Reimund Schwarze:** Can minimum prices assure the quality of professional services?.

2007-06  **Michal Brzoza-Brzezina and Jesus Crespo Cuaresma:** Mr. Wicksell and the global economy: What drives real interest rates?.


2007-03  **Paul Raschky:** The overprotective parent - Bureaucratic agencies and natural hazard management.


2007-01  **Martin Kocher, Matthias Sutter and Florian Wakolbinger:** The impact of naïve advice and observational learning in beauty-contest games.
Gender pairing and bargaining – Beware the same sex!

Abstract
We study the influence of gender and gender pairing on economic decision making in an experimental two-person bargaining game where the other party’s gender is known to both actors. We find that (1) gender per se has no significant effect on behavior, whereas (2) gender pairing systematically affects behavior. In particular, we observe much more competition and retaliation and, thus, lower efficiency when the bargaining partners have the same gender than when they have the opposite gender. These findings are consistent with predictions from psychology. Implications of our results for real-world organizations are discussed.