Open access • Journal Article • DOI:10.1007/S10683-009-9217-9

Gender pairing and bargaining—Beware the same sex! — Source link

Matthias Sutter, Matthias Sutter, Ronald Bosman, Martin G. Kocher ...+2 more authors

Institutions: University of Innsbruck, University of Gothenburg, De Nederlandsche Bank, Ludwig Maximilian University of Munich ...+1 more institutions

Published on: 09 Jun 2009 - Experimental Economics (Springer US)

Topics: Evolutionary psychology

Gender pairing and bargaining—Beware the same sex!

Sutter, M.; Bosman, R.; Kocher, M.G.; van Winden, F.

DOI
10.1007/s10683-009-9217-9
Publication date
2009
Document Version
Author accepted manuscript
Published in
Experimental Economics

Gender pairing and bargaining – Beware the same sex![*]

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

This version: 12 May 2009

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 evolutionary 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: f.a.a.m.vanwinden@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, 2009) or behavior in competitive environments (Gneezy et al., 2003, 2009; Gupta et al., 2005; Niederle and Vesterlund, 2007; Niederle et al., 2008).[1] So, gender[2] 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 evolutionary psychology (Buss, 1999) suggests that gender *pairing* has an effect on (economic) behavior.

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

[1] A comprehensive overview of the literature is provided by Croson and Gneezy (2009).

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.[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 *per se* has no significant effect on behavior, but that gender *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. Section 4 is devoted to the experimental design, while the results are given in section 5. Section 6 relates our findings to insights from evolutionary psychology and concludes the paper with a discussion of the implications of our findings with respect to applied organizational research.

2 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 suggests 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,

[2] We use the expressions „sex“ and „gender“ interchangeably throughout the paper.

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. One way to reconcile these different findings is provided through the results of Andreoni and Vesterlund (2001). They present evidence that the variable gender interacts systematically with the price of altruism, that is the cost of giving away money to the (powerless) responder in the dictator game. When altruism is expensive, women are kinder, but when it is cheap, men are more altruistic. 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

[3] The details of the game and its relevance for economic decision making will be explained in section 3.

responder rejects, both earn nothing. With respect to gender pairing, two studies seem particularly 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.[4,5] 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.[6] With such a design, the effects of gender and gender pairing might easily be confounded with the effects of visual expression or beauty.[7]

[4] 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.

[5] 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.

[6] More precisely, four proposers sat opposite of four responders. Participants were told that they would be paired with one of the (four) opposite players.

[7] 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 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 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. We start out this project with the null hypothesis that gender and gender pairing do not have an effect on bargaining behavior. After presenting the power-to-take game, the experimental design and results we come back to this null hypothesis in the concluding section and relate our findings to insights from evolutionary psychology.

3 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.

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 $E_{take} = E_{resp} = 120$

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.[8] 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.[9]

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.[10]

Our four different treatments (FF, FM, MF, MM) result from a 2×2 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 (instructions are provided as supplementary material online):[11] 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

[8] $t = 100\%$ 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.

[9] Recall that the responder can only destroy his or her own income (E_{resp}), but not that of the take authority (E_{take}).

[10] 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).

[11] The game was framed as neutral as possible, avoiding any suggestive terms like take authority or take rate.

subject in the role of A is a woman, and the subject in the role of B is a man."[12] 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. 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

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

Averaging over all treatments (see the outer right column of Table 1), the take rate equals two thirds of the responder's endowment E_{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_{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

[12] 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.

face females. Average destruction rates (46%) and the relative frequency of destruction (63%) are highest in the MM-treatment, where males interact with males.

Table 2 about here

In order to test for 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.[13] Hence, we cannot reject our null hypothesis of no gender differences *per se*.

In order to test for the effects 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 evidence of higher take rates when the responder is female (FF: 75%) than when the responder is male (FM: 64%) ($p < 0.1$; two-sided Mann-Whitney U-test). Similar effects of gender pairing can be found for male take

[13] 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.

authorities, with higher take rates in MM (70%) than in MF (57%), though the effects fail significance at conventional levels ($p = 0.12$; two-sided Mann-Whitney U-test).[14]

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%) ($p < 0.01$; two-sided U-test). The frequency of destruction is also significantly larger in MM than in FM ($p < 0.01$; two-sided χ^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'.[15] 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 destruction is about 80% larger. Remarkably, under same gender pairing, ten out of 38 decision makers chose $t > 95\%$, whereas this occurs only twice under mixed gender pairing ($p < 0.05$; $\chi^2 = 6.33$; two-sided). Regarding the destruction rates, ten decision makers in the same gender pairing condition chose $d > 95\%$, but only four decision makers in the mixed gender pairing condition ($p < 0.1$; $\chi^2 = 3.15$; two-sided).

[14] We also find significantly higher take rates in FF than in MF ($p < 0.05$; two-sided Mann-Whitney U-test). All other pairwise comparisons yield no significant differences.

[15] Pooling is possible, because take rates, destruction rates and frequencies of destruction do not differ significantly (Mann-Whitney U-test, Kolmogorov-Smirnov-test) between FF and MM (same gender pairing), nor between FM and MF (mixed gender pairing).

Table 3 about here

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.

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 ($p < 0.01$ in FF, $p < 0.05$ in FM, $p < 0.1$ in MF and MM; 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 expectations not having been incentivized, even though Grether (1992) indicates that 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.[16]

Table 4 about here

[16] 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 or female take authority, therefore expecting the same behavior of both male and female take authorities.

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 differs between responders who destroyed something or everything and those who destroyed nothing ($p < 0.1$; $\chi^2 = 3.15$; two-sided test). This suggests that those disappointed by the take rate are more likely to destroy.

6 Discussion and conclusion

Our results on behavior in a bargaining experiment (the power-to-take game) suggest that there are no significant differences between men and women *per se*. However, gender pairing has been identified as 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.

These marked effects of gender pairing can be related to insights from 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 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). 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 the interaction between members of the same sex will be more aggressive or competitive. This is what we indeed have found.

Our findings with respect to gender pairing are also related to recent research that has focused on differences in the competitiveness of men and women. In their seminal study on gender differences in competitive environments Gneezy et al. (2003) have found an interesting effect of gender pairing on the willingness to enter a competition. While women are as likely as men to enter competitive situations (in the form of a tournament) in single-gender competitions, they are less likely to do so in mixed-gender competitions. This implies that women are more competitive in a single-gender environment than in a mixed-gender environment. A similar pattern has been established in this paper in the context of a bargaining game. However, the results for men differ slightly between our study and the one in Gneezy et al. (2003). In the role of take authorities, we also find no significant effects of gender pairing. Yet, male responders significantly differ in their destruction rates and the frequency of destruction depending on gender pairing (even when controlling for take rates). This suggests that men in the role of responders are also more competitive in a same-gender environment than in a mixed-gender one.

When comparing our findings with previous experimental studies on gender differences *per se*, we would like to stress that the existing evidence on the influence of gender in two-person bargaining games, like the ultimatum game or the dictator game, is not fully conclusive (see Camerer, 2003, and Croson and Gneezy, 2009, 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 the case.

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.[17] 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.

References

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

Ayres, I. (1991), Fair driving: Gender and race discrimination in retail car negotiations. *Harvard Law Review* 104: 817-872.

Ayres, I., Siegelman, P. (1995), Race and gender discrimination in bargaining for a new car. *American Economic Review* 85: 304-321.

Ben-Ner, A., Kong, F., Putterman, L. (2004), Share and share alike? Intelligence, socialization, personality, and gender-pairing as determinants of giving. *Journal of Economic Psychology* 25: 581-589

Bolton, G., Katok, I. (1995), An experimental test for gender differences in beneficent behavior. *Economics Letters* 48: 287-292.

Bosman, R., Sutter, M., van Winden, F. (2005), On the role of emotions and real effort in a power-to-take game: Experimental studies. *Journal of Economic Psychology*, 26: 407-429.

Bosman, R., van Winden, F. (2002), Emotional hazard in a power-to-take experiment. *Economic Journal* 112: 147-169.

Bowles, H. R., McGinn, K. L. (2002), When does gender matter in negotiation?. John F. Kennedy School of Government, Harvard University. Working Paper RWP02-036.

Bowles, H. R., Babcock, L., McGinn, K. L. (2005), Constraints and triggers: Situational mechanics of gender in negotiation. *Journal of Personality and Social Psychology* 89: 951-965.

Buss, D. M. (1999), *Evolutionary Psychology. The New Science of the Mind.* Allyn and Bacon, Boston.

Camerer, C. F. (2003), *Behavioral Game Theory. Experiments in Strategic Interaction*. Princeton University Press.

[17] 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 misleading conclusions.

Camerer, C., Hogarth, R. M. (1999): The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* 19: 7-42.

Carpenter, J., Verhoogen, E., Burks, S. (2005). The effects of stakes in distribution experiments. *Economics Letters* 86: 393-398.

Craver, C., Barnes, D. (1999), Gender, risk taking, and negotiation performance. *Michigan Journal on Gender and Law* 5: 299-352.

Croson, R., Buchan, N. (1999), Gender and culture: International experimental evidence from trust games. *American Economic Review*, *Papers and Proceedings* 89: 386-391.

Croson, R., Gneezy, U. (2009), Gender differences in preferences. *Journal of Economic Literature*, forthcoming.

Dufwenberg, M., Muren, A. (2006), Gender composition in teams, *Journal of Economic Behavior and Organization* 61: 50-54.

Eckel, C. C., Grossman, P. J. (1998), Are women less selfish than men? Evidence from dictator experiments. *Economic Journal* 108: 726-735.

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

Eckel, C. C., Grossman, P. J. (2008), Differences in the economic decision making of men and women: Experimental evidence. In. Plott, C., Smith, V.L. (eds.), *Handbook of Experimental Economics Results*. Vol. 1, North Holland, Amsterdam: 509-519.

Fehr, E., Fischbacher, U., von Rosenbladt, B., Schupp, J., Wagner, G. G. (2003), A nationwide laboratory. Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. Institute for Empirical Research in Economics. Working Paper 141. University of Zurich.

Fehr, E., Naef, M. and Schmidt, K. (2006), The role of equality, efficiency, and Rawlsian motives in social preferences: A reply to Engelmann and Strobel, *American Economic Review* 96: 1912-1917.

Fischbacher, U. (2007), z-Tree: Zurich toolbox for readymade economic experiments. *Experimental Economics* 10: 171-178.

Frey, B., Bohnet, I. (1995), Institutions affect fairness: Experimental investigations. *Journal of Institutional and Theoretical Economics* 151: 286-303.

Frey, B. Bohnet, I. (1999a), Social distance and other-regarding behavior in dictator games: Comment. *American Economic Review* 89: 335-339.

Frey, B., Bohnet, I. (1999b), The sound of silence in prisoner's dilemma and dictator games. *Journal of Economic Behavior and Organization* 38: 43-57.

Ginther, D. K., Kahn, S. (2004), Women in economics: Moving up or falling off the academic career ladder?. *Journal of Economic Perspectives* 18(3): 193-214.

Ginther, D. K., Hayes, K. J. (2003), Gender differences in salary and promotion for faculty in the humanities. *Journal of Human Resources* 38: 34-73.

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

Gneezy, U., Leonard, K. L., List, J. A. (2009), Gender differences in competition. Evidence from a matrilineal and a patriarchal society. *Econometrica*, forthcoming.

Grether, D. M. (1992), Testing Bayes rule and the representative heuristic: Some experimental results. *Journal of Economic Behavior and Organization* 17: 623-638.

Gupta, N.D., Poulsen, A., Villeval, M. (2005), Do (wo)men prefer (non-)competitive jobs? GATE Working Paper No. 05-12 and IZA Discussion Paper No. 1833.

Holm, H. J. (2000), Gender-based focal points. *Games and Economic Behavior* 32: 292-314.

Loewenstein, G. (2000), Emotions in economic theory and economic behavior. *American Economic Review, Papers and Proceedings*, 90: 426-432.

Niederle, M., Segal, C., Vesterlund, L. (2008), How costly is diversity? Affirmative action in light of gender differences in competitiveness. Working Paper.

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

Robertson, R. E. (2001), *Women in Management: Analysis of Selected Data from the Current Population Survey*. (Report to Congressional Requesters No. GAO Report 02-156). U.S. General Accounting Office, Washington, D.C.

Schweitzer, M., Solnick, S. (1999): The influence of physical attractiveness and gender on ultimatum game decisions. *Organizational Behavior and Human Decision Processes* 79: 199-215.

Sokoloff, N. J. (1992), *Black Women and White Women in the Professions: Occupational Segregation by Race and Gender, 1960-1980.* Routledge, New York.

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

Stuhlmacher, A. F., Walter, A. E. (1999), Gender differences in negotiation outcome: A meta-analysis. *Personnel Psychology* 52: 653-677.

Sutter, M., Kocher, M.G. (2007), Trust and trustworthiness across different age groups. *Games and Economic Behavior* 59: 364-382.

Trivers, R. L. (1972), Parental investment and sexual selection. In: Campbell, B. (ed.), *Sexual Selection and the Descent of Man: 1871-1971*. Aldine, Chicago: 136-179.

Walters, A. E., Stuhlmacher, A. F., Meyer, L. L. (1998), Gender and negotiator competitiveness: A meta-analysis. *Organizational Behavior and Human Decision Processes* 76: 1-29.

Watson, C. (1994), Gender versus power as a predictor of negotiation behavior and outcomes. *Negotiation Journal* 10: 117-127.

Tables and Figures

Table 1. Decisions

		treatment*				
		FF	FM	MF	MM	overall
take rate (%)	average	75.42[a,b]	63.89[a]	57.16[b]	70.21	66.67
	(standard deviation)	(20.13)	(15.72)	(25.24)	(25.56)	(22.65)
destruction rate (%)	average	36.63[a]	13.42[a,c]	24.32	45.84[c]	30.05
	(standard deviation)	(43.62)	(31.71)	(35.29)	(41.94)	(39.64)
frequency of destruction (%)	average	52.63[a]	21.05[a,c]	42.11	63.16[c]	44.74
profit take authority[#]	ECU	230.54	243.76	228.72	218.36	230.34
profit responder[#]	ECU	85.50	100.31	102.10	86.63	93.64

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

[#] including show up fee of 60 ECU.

[a] significantly different in pairwise comparison at $p < 0.1$ (two-sided-tests)

[b,c] significantly different in pairwise comparison at $p < 0.05$ (two-sided-tests)

$N = 19$ for each single treatment.

Table 2. Decisions grouped by gender and gender pairing

	gender			*gender pairing*		
	females	males	significance	same	mixed	significance
take rate (%)	69.66	63.68	n.s.	72.82	60.53	$p < 0.05$ (two-sided U-test)
destruction rate (%)	30.47	29.63	n.s.	41.24	18.87	$p < 0.05$ (two-sided U-test)
frequency of destruction (%)	47.37	42.11	n.s.	57.89	31.58	$p < 0.05$ (two-sided χ^2-test)

n.s. not significant.

Table 3. Take rates and destruction rates

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

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

	treatment				
	FF	FM	MF	MM	overall
expected take rate in % (average)	41.58	44.42	39.21	50.53	43.93
(standard deviation)	(33.08)	(29.42)	(26.84)	(24.26)	(28.33)
actual take rate in %	75.42	63.89	57.16	70.21	66.67
expected destruction rate in % (average)	23.95	16.58	5.26	19.58	16.34
(standard deviation)	(33.69)	(29.06)	(9.79)	(32.37)	(28.25)
actual destruction rate in %	36.63	13.42	24.32	45.84	30.05
expected frequency of destruction in %	52.63	36.84	31.58	47.37	42.11
actual frequency of destruction in %	52.63	21.05	42.11	63.16	44.74

Figure 1. Actual vs. expected take rate and destruction (N = 76)