Experimental evidence of the importance of gender pairing in bargaining*

Matthias Sutter[#], Ronald Bosman[+], Martin Kocher[#], and Frans van Winden[&]

This version: 3rd September 2003

Abstract

We study the influence of gender on economic decision making in a two-person bargaining game. By testing hypotheses derived from evolutionary psychology and social role theory 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. Implications for real-world organizations are discussed.

Key words: gender pairing, bargaining, principal-agent relationship, psychology, experiment.
JEL-classification: A12, C72, C91, C92.

[#]University of Innsbruck
Institute of Public Economics
Universitätsstraße 15
A-6020 Innsbruck
Austria
E-mail1: martin.kocher@uibk.ac.at
E-mail2: matthias.sutter@uibk.ac.at
Sutter is corresponding author.

[+] Monetary and Economic Policy Department
De Nederlandsche Bank
Westeinde 1
1017 ZN Amsterdam
The Netherlands
E-mail: r.a.j.bosman@dnb.nl

[&] CREED/Department of Economics
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam
The Netherlands
E-mail: fvwinden@fee.uva.nl

* This research was supported by a grant from the Austrian National Bank (Jubiläumsfonds-Projekt Nr. 8487).

1 Introduction

Many 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 hirings and firings (Ginther and Kahn, 2003) or the efficiency of workgroups (Levine and Moreland, 1998). Apparently, gender has an impact on the functioning of organizations. Some discrepancies, in particular those with respect to salaries, promotions and hirings, have been scrutinized for possible differences in the bargaining behavior of men and women (Watson, 1994). However, the evidence on gender differences in bargaining is rather mixed, as we will show in greater detail in section 2. One possible reason for the mixed, and frequently insignificant, results on gender and bargaining may be the fact that gender *pairing* is often not accounted for in such studies. A broad literature in social psychology and, in particular, evolutionary psychology suggests that gender pairing should have a significant impact on economically relevant behavior, as will be discussed in section 3.

In this paper, we present an experimental study examining the effects of gender and gender pairing under controlled laboratory conditions. Our vehicle of research is the power-to-take game (see Bosman and van Winden, 2002, Bosman et al., 2003), which is a two-person bargaining game that relates to some important economic situations, such as principle-agent relationships.[1] Based on theoretical predictions from social and evolutionary psychology, we assess the importance of gender pairing by looking at the four possible constellations with regard to gender in a 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

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

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 in bargaining. Section 3 describes the power-to-take game and presents some hypotheses on the effects of gender and gender pairing based on the 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.

2 Gender and bargaining

Before experimental economics got interested in the effects of gender and gender pairing on economic behavior, several field studies had already investigated these effects. Two of the most prominent ones in economics are Ayres (1991) and Ayres and Siegelman (1995). They find that gender *pairing* is a significant variable in negotiations for the purchase of a new car. In particular, women receive worse deals from women than from men. With respect to the influence of gender on negotiation behavior, for instance on salaries, the evidence is mixed, however (see Menkel-Meadow, 2000, for a survey). The meta-analyses of the influence of gender on negotiation 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), on the contrary, claim that there are no statistically significant differences in negotiation outcomes and performances between men and women. A possible explanation for the mixed results might be the lack of control and standardization that is inevitable when working with field data. Therefore, one might hope to get less ambiguous results from controlled experiments.

The experimental dictator game provides a good starting point for studying the effects of gender on bargaining behavior.[2] 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.[3] To date, the evidence on gender effects in the dictator game is ambiguous, though. Whereas Bolton and Katok (1995) or Frey and Bohnet (1995) find no evidence for gender differences, Eckel and Grossman (1998) report women to be significantly less selfish than men. Recently, Andreoni and Vesterlund (2001) have found 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. They find that when altruism is expensive, women are kinder, but when it is cheap, men are more altruistic.

In order to study real bargaining behavior, 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. 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. One possible reason for the different findings might be differences in the experimental procedure. In Solnick's (2001) study participants sat in

[2] Camerer (2003) provides a brief overview of experimental studies on the influence of gender on bargaining behavior.

[3] Note, however, that Schubert et al. (1999), for instance, find no gender differences in risk attitudes.

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.[4] With such a design, the effects of gender pairing might easily be confounded with the effects of visual expression or beauty.[5]

Summarizing, there is mixed evidence on the role of gender and gender pairing in bargaining, both from field studies as well as experimental studies. It is not easy to explain why the evidence is so mixed. Studies differ in many important ways, like the way gender pairing is controlled for, the way in which the bargaining game is implemented, or the way in which gender is revealed. Our study will carefully control for gender as well as gender pairing and provide some testable hypotheses, which will be discussed in the following section.

3 The power-to-take game and hypotheses from evolutionary and social 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

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

[5] Schweitzer and Solnick (1999), for instance, have found in an ultimatum game that there is something like a beauty premium, meaning that more attractive people are offered more.

authority the payoff is thus given by $E_{take} + t(1\text{-}d)E_{resp}$. For the responder, the payoff equals $(1\text{-}t)(1\text{-}d)E_{resp}$.

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. Hostile take-overs may have similar characteristics with one party claiming another party's assets and the other party being able to make the take-over very costly by influencing the value of its assets. The power-to-take game can also 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 her or 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.

3.2 Hypotheses

3.2.1 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 (intrasexual 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 would seem to predict that interaction between members of the same sex will be more aggressive or competitive. This is stated in our first hypothesis:

Hypothesis 1: *Take rates and destruction rates in the power-to-take game will be higher under same gender pairing than under mixed gender pairing.*

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 *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 (see Macoby and Jacklin, 1974; Hyde, 1986; 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, we cannot derive from evolutionary psychology any hypothesis on the influence of gender *per se* on bargaining. Social psychology, however, provides further guidance.

3.2.2 Social psychology

Social psychology offers two prominent alternative theories for explaining gender differences in social behavior: social role theory and status characteristics theory (see Carli and Eagly, 1999, and Eagly et al., 2000). According to these theories the sexes differ with respect to their social position, in particular regarding their status and social roles. Gender differences are not seen as stable and constant across individuals, as evolutionary theory assumes, but as being dependent on particular (sub)cultures, historical periods, and even individual life histories. In addition, gender differences are taken to depend on the salience of gender norms and the influence of social roles that compete with gender roles.

According to status characteristics theory, people form expectations about how other people will behave on the basis of status and perceived competence. Generally, people tend to have more status if they are male rather than female, white rather than black, more educated rather than less educated, or physically attractive rather than unattractive. Because in Western societies more respect, honor and importance are attached to men than women, gender can be seen as a proxy for status.

Social role theory assumes that gender differences are the result of differences in the division of labor between men and women. Women are more likely than men to be primary care takers, emphasizing behaviors such as being subordinate, sensitive, more cooperative, etc. In other words, women show more 'communal' behavior. Men, on the other hand, are more likely to be involved in occupational roles. In addition, the occupational roles of men are different than those of women, requiring behaviors such as self-assertion, achievement, competition, and even aggression. The behavior of men is therefore referred to as being 'agentic'.

According to Carli and Eagly (1999, p. 208) "both social role theory and status characteristics theory predict that men should exert greater influence over others than women do." This prediction seems to imply for the power-to-take game that male take authorities will choose higher take rates than female take authorities, and that male responders will choose higher destruction rates than female responders.[6] This is summarized in our second hypothesis:

Hypothesis 2: *Compared to women, men will choose higher take rates and destruction rates in the power-to-take game.*

4 Experimental design

The experiment was computerized with the help of z-Tree (Fischbacher, 1999). Sessions were run in May and June 2001 at the University of Innsbruck. For each treatment, we ran 2 sessions with 19 pairs, involving 38 subjects.[7] About 75% of our 152 participants were

[6] Note that when males bargain with males, it might be the case that male take authorities anticipate correctly that male responders have higher destruction rates – controlling for the take rate – then female responders. This expectation might, then, lead male take authorities to reduce their (intended) take rates when facing male responders, possibly to the level of female take authorities. Consequently, such expectations could imply that there are no gender differences observable at all, contrary to our Hypothesis 2. Yet, it is unclear whether male (female) take authorities are able to anticipate the effect of gender on the responder's behavior correctly. The well-documented hot-cold empathy gap (see Loewenstein, 2000), for instance, documents that people are bad in predicting behavior in a state they are not currently in themselves.

[7] All sessions were conducted by the same experimenter.

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 Austrian Schilling (about $12).

At the beginning of the experiment, participants received a show up fee of 60 Austrian Schilling and 120 Austrian Schilling as initial endowment (E_{take}, E_{resp}). 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]

[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).

We have four different treatments, resulting from a 2x2 matrix determined by the take authority's and the responder's gender. Subjects were informed about the gender of both roles in the instructions in the following way (see appendix A2).[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 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.

5 Results

We divide our data analysis in 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.

Averaging over all treatments, 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. $d > 0$). Looking at

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

[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.

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 play against males. 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 ATS vs. 94 ATS).

In order to test our first hypothesis, we control for gender when comparing decisions in treatments with same gender pairing, respectivly mixed gender pairing. 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[13]). 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).[14]

Table 1. Decisions

		treatment*				
		FF	FM	MF	MM	overall
take rate (%)	average	75.42	63.89	57.16	70.21	66.67
	(standard deviation)	(20.13)	(15.72)	(25.24)	(25.56)	(22.65)
destruction rate (%)	average	36.63	13.42	24.32	45.84	30.05
	(standard deviation)	(43.62)	(31.71)	(35.29)	(41.94)	(39.64)
frequency of destruction (%)	average	52.63	21.05	42.11	63.16	44.74
profit take authority#	ATS	230.54	243.76	228.72	218.36	230.34
profit responder#	ATS	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 ATS.

$N = 19$ for each single treatment.

[13] Since both of our hypotheses provide a directional prediction concerning the effects of gender and gender pairing on take rates or destruction rates, we can apply one-sided tests in the statistical analysis.

[14] We also find significantly higher take rates in FF than in MF ($p = 0.011$; one-sided Mann-Whitney U-test). All other pairwise comparisons yield no significant differences.

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 χ^2-test). For female responders, gender pairing has no significant effect on destruction rates and the frequency of destruction.[15]

Table 2. Decisions grouped by gender pairing and gender

	gender pairing			*gender*		
	same	mixed	significance (same vs. mixed)	females	males	significance
take rate (%)	72.82	60.53	$p < 0.01$ (one-sided U-test)	69.66	63.68	n.s.
destruction rate (%)	41.24	18.87	$p < 0.01$ (one-sided U-test)	30.47	29.63	n.s.
frequency of destruction (%)	57.89	31.58	$p < 0.02$ (one-sided χ^2-test)	47.37	42.11	n.s.

n.s. not significant.

Another way to show the effects of gender pairing is to pool treatments by gender pairing, as is done on the left-hand side of Table 2. Treatments FF and MM are pooled to 'same gender pairing', and FM and MF to 'mixed gender pairing'.[16] 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 column stating significance levels in Table 2. Take rates are about 20% higher when subjects face the same gender than if 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

[15] Note, however, that destruction rates and the frequency of destruction are significantly higher in FF than in FM (p < 0.05; one-sided Mann-Whitney U-test and χ^2-test, respectively).

[16] 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 ($p < 0.02$; $\chi^2 = 6.33$; 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 ($p < 0.05$; $\chi^2 = 3.15$; one-sided).

Table 3. Take rates and destruction rates

take rate	same gender destruction rate (average)	N	mixed gender destruction rate (average)	N
0-10%	0	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 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.[17] We summarize our evidence in

[17] In the interval [0%, 10%], we cannot compare the two conditions because for the same gender pairing we have no observation in this interval.

Result 1: *Take rates and destruction rates are higher under same gender pairing than under mixed gender pairing.*

On the right-hand side of Table 2 we aggregate treatments by the gender of the decision maker. For example, female take rates consider treatments FF and FM, while female destruction rates regard FF and MF. We find no significant differences between females and males with respect to take rates, destruction rates and frequencies of destruction.[18] Even when we control for gender pairing, we do not find different behavior of men in women (comparing either FF with MM or FM with MF). Hence, we do not find support for Hypothesis 2.

Result 2: *There are no differences in behavior between females and males, even when controlling for gender pairing. Take rates and destruction rates do not depend on gender.*

5.2 Expectations

5.2.1 Expected versus actual decisions

Table 4 reports expected take and destruction rates and compares them to actual decisions. Overall, responders expect a take rate of about 44%, which falls about 22 percentage points short of the average actual take rate. The expected destruction rate (16%) falls about 14 percentage points short of the overall average destruction rate.

[18] All *p*-values are larger than 0.1 for one-sided tests.

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

Expected take rates are significantly smaller than the actual ones in each single treatment ($p < 0.01$ 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. Apparently responders do not take into account that take authorities of the same gender behave more aggressively, that is choose a higher take rate on average.

Expected destruction rates, on the other hand, do show a gender effect. 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, nor between men and women.[19] However, expected destruction rates are significantly higher in the same gender pairing-condition than with mixed gender, which is quite intuitive since the take authorities in case of same gender pairing had chosen a higher take rate.

Table 5. Expected versus actual decisions – grouped by gender and gender pairing

	gender pairing			gender		
	same	mixed	significance	females	males	significance
expected take rate in % (average)	46.05	41.82	n.s.	40.39	47.47	n.s.
(standard deviation)	(28.97)	(27.90)		(29.74)	(26.78)	
actual take rate in %	72.82	60.53	$p < 0.01$ (one-sided U-test)	69.66	63.68	n.s.
expected destruction rate in % (average)	21.76	10.92	$p = 0.05$ (one-sided U-test)	20.26	12.42	n.s.
(standard deviation)	(32.66)	(22.14)		(31.26)	(24.68)	
actual destruction rate in %	41.24	18.87	$p < 0.01$ (one-sided U-test)	30.47	29.63	n.s.
expected frequency of destruction in %	50.00	34.21	n.s.	44.74	39.47	n.s.
actual frequency of destruction in %	57.89	31.58	$p < 0.02$ (one-sided χ^2-test)	47.37	42.11	n.s.

n.s. not significant.

5.2.2 The influence of expected take rates on actual destruction rates

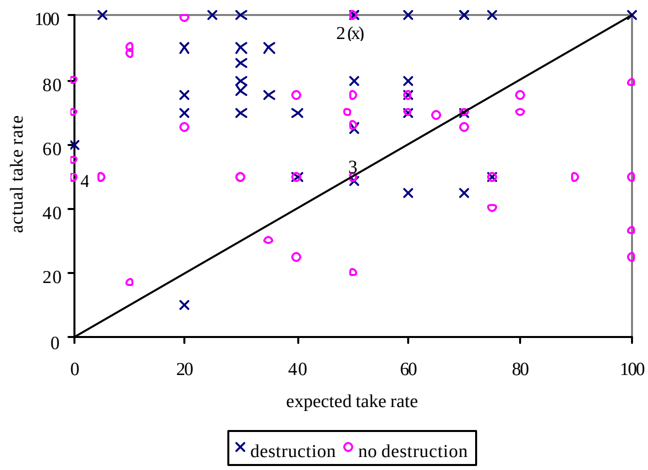
Figure 1 plots individual data on expected take rates (on the horizontal axis) versus actual take rates (on the vertical axis). Points above (below) the diagonal indicate that expectations were lower (higher) than actual decisions, and, thus, too optimistic (pessimistic). We have marked those responders who destroyed parts or all of their endowment by a cross.[20]

[19] FF versus MF comes closest to being significant ($p = 0.11$; two-sided Mann-Whitney U-test). All other pairwise comparisons have $p > 0.25$ (two-sided Mann-Whitney U-tests).

[20] It is noteworthy that 8 responders (4 in each gender pairing condition) expected a take rate of zero, but that 7 of them did not destroy anything, even though actual take rates ranged from 50% to 80%. It is possible that these subjects wanted to express to have no expectation at all. In the experiment by Bosman and van Winden (2002),

The frequency of points lying above or below the diagonal is significantly different between responders who destroyed something or everything and those who destroyed nothing ($p < 0.05$; $\chi^2 = 3.15$; one-sided test).

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



To study the role of expectations for actual destruction rates more thoroughly, we have estimated a censored tobit regression model. The dependent variable is the destruction rate, with left and right censoring at $d = 0\%$ and $d = 100\%$, respectively. As independent variables, we take the difference between the actual and expected take rate and a dummy variable for gender pairing (1 for same gender; 0 for mixed gender). Results are shown in Table 6.

The difference between the actual and expected take rate turns out to have a significantly positive influence on the destruction rate d. Likewise, gender pairing has a significant impact on the destruction rate, with a higher probability of destruction when the take authority and responder have the same gender. The latter result provides further evidence

where responders were given the option to indicate not to have any expectation, 17 out of 39 responders chose this option. However, in our experiment, responders had to type in a figure for expected take rates.

that destruction rates are significantly higher under same gender pairing than under mixed gender pairing.

Table 6– Determinants of the destruction rate (censored tobit regression)

Independent variable	Coefficient	Z-value	Probability
constant	-67.35	-2.40	0.016
actual – expected take rate	1.01	2.33	0.019
same gender pairing (1 yes, 0 no)	61.17	2.11	0.035
dependent variable: destruction rate			
left censored observations	42		
right censored observations	12		
N	76		
R^2	0.15		
Log likelihood	-169.56		

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 intermale and interfemale bargaining. Furthermore, we find no significant differences in behavior between men and women, even when controlling for gender pairing. Overall, our results are in line with predictions derived from evolutionary psychology, but are largely at odds with hypothesized differences between men and women derived from social role theory or status characteristics theory.

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 far from conclusive (see Camerer, 2003, for a survey). 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 evidence for no gender differences. 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 there actually are.

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.[21] 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 clear 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 probability that scarce resources will be wasted.

[21] 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

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

Archer, J. (1996), Sex differences in social behavior, *American Psychologist* 51: 909-917.

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.

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. (2003), On the role of emotions and real effort in a power-to-take game: Experimental studies. *Journal of Economic Psychology*, forthcoming.

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

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.

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

Campbell, A. (1995), A few good men: Evolutionary psychology and female adolescent aggression. *Ethology and Sociobiology* 16: 99-123.

Carli, L.L., Eagly, A.H. (1999). Gender effects on social influence and emergent leadership, in: Powell, N., Handbook of gender and work, Sagem Publications, London.

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

Eagly, A.H., Wood, W., Diekman, A.B. (2000). Social role theory of sex differences and similarities: A current approach, in: Ecks, T., Trautner, H., Developmental social psychology of gender, Lawrence Erlbaum & Associates Publishers, New Jersey.

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.

Fischbacher, U. (1999), Z-tree: Zurich toolbox for readymade economic experiments. Working paper No. 21, Institute for Empirical Research in Economics, University of Zurich.

Fischer, A. H., Rodriguez Mosquera, P. M. (2001), What concerns men? Women or other men? A critical appraisal of the evolutionary theory of sex differences in aggression. *Psychology, Evolution and Gender* 3: 5-25.

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

Ginther, D. K., Kahn, S. (2003), Women in economics: Moving up or falling off the academic career ladder?. *Journal of Economic Perspectives*, forthcoming.

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.

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

Hyde, J. S. (1984), How large are gender differences in aggression. *Developmental Psychology* 20: 722-736.

Hyde, J. S. (1986), Gender differences in aggression. in: Hyde, J. S., Linn, M. C. (eds.), *The Psychology of Gender: Advances through Meta-Analysis*. Baltimore, Johns Hopkins University Press: 51-66.

Kinney, T. A., Smith, B. A., Donzella, B. (2001), The influence of sex, gender, self-discrepancies, and self-awareness on anger and verbal aggressiveness among U.S. college students. *Journal of Social Psychology* 141: 245-275.

Knight, G. P., Fabes, R. A., Higgins, D. A. (1996), Concerns about drawing causal inferences from meta-analyses: An example in the study of gender differences in aggression. *Psychological Bulletin* 119: 410-421.

Levine, J. M., Moreland, R. L. (1998), Small groups. in: Gilbert, D. T. et al. (eds.) *The handbook of social psychology.* 4th edition, Vol. 2: 415-469.

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

Macoby, E. E., Jacklin, C. N. (1974), *The psychology of sex differences.* Stanford University Press.

Menkel-Meadow, C. (2000), Teaching about gender and negotiation: Sex, truths, and videotape. *Negotiation Journal* 16: 357-375.

Ramirez, J. M., Fujihara, T., van Goozen, S. (2001), Cultural and gender differences in anger and aggression: A comparison between Japanese, Dutch, and Spanish students. *Journal of Social Psychology* 141: 119-121.

Schubert, R., Brown, M., Gysler, M., Brachinger, H. W. (1999), Financial decision-making: Are women really more risk-averse?. *American Economic Review, Papers and Proceedings*, 89: 381-385.

Schweitzer, M. and S. Solnick (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.

Trivers, R. L. (1972), Parental investment and sexual selection. in: Campbell, B. (ed.), *Sexual selection and the descent of man: 1871-1971.* Aldine, Chicago: pp. 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.

Appendix

A1: Individual raw data

Table A1. Individual data

female vs. female

pair	take rate	destruction rate	expected take rate	expected destruction rate
1	45	3	70	0
2	50	0	30	0
3	50	0	0	30
4	50	0	5	0
5	50	0	50	65
6	60	50	0	0
7	70	0	0	0
8	70	30	70	0
9	70	0	80	40
10	80	65	60	20
11	80	0	0	10
12	80	98	50	100
13	88	0	10	60
14	90	0	10	0
15	100	100	50	10
16	100	100	75	100
17	100	100	100	0
18	100	50	70	0
19	100	100	60	20

female vs. male

pair	take rate	destruction rate	expected take rate	expected destruction rate
1	30	0	35	10
2	50	25	40	0
3	50	0	30	0
4	50	0	5	100
5	50	0	75	0
6	50	0	0	70
7	50	0	90	0
8	55	0	0	0
9	65	0	20	35
10	65	0	70	40
11	70	0	49	0
12	70	0	60	0
13	75	0	80	0
14	75	0	60	0
15	75	0	40	0
16	79	0	100	50
17	80	100	30	0
18	85	100	30	10
19	90	30	30	0

male vs. female

pair	take rate	destruction rate	expected take rate	expected destruction rate
1	10	50	20	5
2	17	0	10	0
3	25	0	40	0
4	40	0	75	0
5	45	60	60	35
6	50	0	100	20
7	50	0	40	0
8	50	0	50	0
9	50	0	0	0
10	50	0	50	0
11	50	0	0	0
12	65	65	50	10
13	70	27	60	0
14	75	50	20	10
15	75	0	50	0
16	75	100	60	0
17	90	10	35	20
18	99	0	20	0
19	100	100	5	0

male vs. male

pair	take rate	destruction rate	expected take rate	expected destruction rate
1	20	0	50	0
2	25	0	100	0
3	33	0	100	0
4	49	45	50	10
5	50	62	75	20
6	66	0	50	0
7	69	0	65	0
8	70	0	70	0
9	70	100	30	0
10	70	70	40	50
11	70	35	20	10
12	75	30	35	0
13	77	50	30	12
14	90	99	20	50
15	100	100	25	0
16	100	0	50	100
17	100	80	30	0
18	100	100	50	100
19	100	100	70	20

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 Austrian Schillings (ATS). You will receive the show up fee irrespective of your decisions in the experiment.

Initial endowment
Each participant receives an endowment of 120 ATS.

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 ATS.
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 ATS (60% of 120 ATS).
The total payoff for A at the end of the experiment is equal to 252 ATS (namely, the show-up fee of 60 ATS plus the initial endowment of 120 ATS plus the transfer of 72 ATS).
The total payoff for B at the end of the experiment is equal to 108 ATS (namely, the show-up fee of 60 ATS plus the endowment of 120 ATS minus the transfer of 72 ATS).

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 ATS (namely, 60% of the remaining endowment of participant B after phase II, which is 60% of 60 ATS).
The total payoff for A at the end of the experiment is equal to 216 ATS (namely, the show-up fee of 60 ATS plus the endowment of 120 ATS plus the transfer of 36 ATS) and for participant B 84 ATS (namely, the show-up fee of 60 ATS plus the remaining endowment of 60 ATS after destruction minus the transfer of 36 ATS).

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.