Chapter 57

DIFFERENCES IN THE ECONOMIC DECISIONS OF MEN AND WOMEN: EXPERIMENTAL EVIDENCE

CATHERINE C. ECKEL

School of Social Sciences, University of Texas at Dallas, Richardson, TX 75080, USA

PHILIP J. GROSSMAN

Department of Economics, Saint Cloud State University, St. Cloud, MN 56301, USA

Abstract

This chapter reviews the results from public goods, ultimatum, and dictator experiments for evidence of systematic differences in the behavior of men and women. While the results do not offer consistent evidence of behavioral differences between men and women, there are some intriguing patterns in the data. No significant evidence of systematic differences in the play of men and women is evident in those settings where subjects are exposed to risk. In those settings where risk is absent, systematic differences are revealed. This finding is conditioned by the level of risk.

Do individual decisions vary systematically by the sex of the decision maker? This question has generated considerable debate. Evidence from the other social and behavioral sciences finds substantial differences between the sexes, and tends to support the stereotypes of the more socially-oriented (selfless) woman and the more individually-oriented (selfish) man.[1]

Dating back at least to Rapoport and Chammah (1965), experimenters have tested in the laboratory for differences in behavior between men and women in situations involving salient monetary incentives. The study by Rapoport and Chammah is an early example in a long series of studies that employ variations of the prisoner's dilemma game to test for such differences. The results have been mixed, with some studies finding women more selfless, others finding men more selfless, and still others finding no difference.[2] More recently, researchers have turned to public goods, ultimatum, and dictator experiments, which we survey here. A critical difference among these experimental designs is the extent to which subjects are exposed to a risk of financial loss, exploitation, or the judgment of others. If women and men respond differently to these risks, this may confound underlying differences in the tendency to cooperate, muddying experimental results in some settings and producing behavioral differences in others.

---

[1] See Eckel and Grossman (1998) for a brief review.
[2] Ibid.

1. Public Goods Experiments

Table 1 reports the results from seven recent public goods experiments. Like the prisoner's dilemma experiments before them, these experiments offer no clear evidence of a systematic difference between men and women. Brown-Kruse and Hummels (1993), Sell and Wilson (1991), and Sell, Griffith, and Wilson (1993), find that women contribute less to the public good than do men. Nowell and Tinkler (1994), and Seguino, Stevens, and Lutz (1996) report significantly higher contributions by women than by men.[3] Cadsby and Maynes (1998) and Sell (1997) report mixed results.[4] In same gender groups, women contribute more than men, though this difference is not significant. When men are the minority in mixed gender groups, they contribute at a significantly higher rate than do women when they are the minority in mixed gender groups.

Methodological differences might explain the varying results. In Brown-Kruse and Hummels (1993), Nowell and Tinkler (1994), Sell and Wilson (1991), Sell, Griffith, and Wilson (1993), and Cadsby and Maynes (1998) the game was repeated (6, 13, 10, 18, 18, and 25 periods, respectively), while Seguino, Stevens, and Lutz (1996) was one-shot. Brown-Kruse and Hummels (1993) constrain their subjects to make all-or-nothing allocations of $1 to the public good. The other studies allow subjects to contribute any number of tokens from zero to their full allocation.[5] Cadsby and Maynes (1998) require contributions reach a threshold level (25 percent of all tokens) before a return is earned. Contributions to the public good involve a kind of risk; other group members may not reciprocate. Subjects risk being "exploited," and may feel as if they have been "played for a sucker" (Ingram and Berger, 1977). While a subject may be inclined to contribute something to the public good, the all-or-nothing design of Brown-Kruse and Hummels (1993) precludes such hedging. By requiring subjects to contribute everything, the design increases the vulnerability of subjects to exploitation and the perceived potential return may not justify this exposure. If women are more risk averse or more afraid of the judgment of others (of looking bad by being taken advantage of), they may be less willing to contribute the full dollar.

A second important design difference making comparison of results difficult is the ability of subjects to monitor the behavior of others. The Sell, Griffith, and Wilson (1993) study provides subjects no such information until after all decisions are made.

---

[3] The difference Sell, Griffith, and Wilson (1993) report is not significant. However, Sell, Griffith, and Wilson (1993) also report results from experiments in which the resource subjects earned was "time with an expert." Results from these sessions reveal the same, but now significant, pattern of behavior. On average, men contributed 61.1% of their tokens to the public good, while women contributed 48.7%. They suggest that this may be because men and women value some resources differently.

[4] Cadsby and Maynes (1998) also replicate Brown-Kruse and Hummels's (1993) procedures and report no significant difference.

[5] The number of tokens given to subjects each period varies: Seguino, Stevens, and Lutz (1996) gave their subjects an allocation of 5 tokens, Cadsby and Maynes (1998) gave 10, Sell and Wilson (1991) and Sell, Griffith, and Wilson (1993) gave 30, and Nowell and Tinkler (1994) gave 62.

Table 1
Public goods experiments. Percent of endowment contributed to public good production. Voluntary contribution mechanism experiments

Period[a]	Subjects[c]	Group size[d]	All-male groups				All-female groups				Mixed groups							
												Male subjects			Female subjects		
			# of groups	First (%)	Last (%)	Average (all periods) (%)	# of groups	First (%)	Last (%)	Average (all periods) (%)	# of groups	First (%)	Last (%)	Average (all periods) (%)	First (%)	Last (%)	Average (all periods) (%)
Brown-Kruse and Hummels[b]	64	4	8	93.8[e]	46.9	67.7[e]	8	71.9[e]	40.6	56.3[e]	NA	NA	NA	NA	NA	NA	NA
Nowell and Tinkler	64	4	5	46.1	14.6	29.6[f]	6	50.5	29.5	37.4[f]	5	NA	NA	NA	NA	NA	NA
Cadsby and Maynes	160	10	8	29.3	NR	19.2	8	37.6	NR	18.9	NA	NA	NA	NA	NA	NA	NA
Seguino, Stevens, and Lutz	139	5–52	NA	NA	NA	NA	NA	NA	NA	NA	6	49.2[g]	NA	NA	65.6[g]	NA	NA
Sell and Wilson[b]	92	4	NA	NR	NR	NR	NA	NR	NR	NR		NR	NR	50.6[e]	NR	NR	36.5[e]
Sell, Griffith, and Wilson	99	4		NR	NR	60.4		NR	NR	47.1		NR	NR	63.0	NR	NR	48.0

Table 1

(*continued*)

Period[a]	Subjects[c]	Group size[d]	All-male groups				All-female groups				# of groups	Mixed groups					
												Male subjects			Female subjects		
			# of groups	First (%)	Last (%)	Average (all periods) (%)	# of groups	First (%)	Last (%)	Average (all periods) (%)		First (%)	Last (%)	Average (all periods) (%)	First (%)	Last (%)	Average (all periods) (%)
Sell	244	4		NR	NR	43.0		NR	NR	51.8		NR	NR	57.1	NR	NR	46.4

Note. NA – not applicable. NR – not reported.

[a]The number of periods per session varies in each paper: Brown-Kruse and Hummels – six periods; Nowell and Tinkler – thirteen periods; Cadsby and Maynes – twenty-five periods; Seguino, Stevens, and Lutz – one period; Sell, Sell and Wilson – ten periods; and Sell, Griffith, and Wilson – eighteen periods. The group fund multiplier also varied across the studies: Brown-Kruse and Hummels – for one-half of the sessions, the multiplier was 1.2 for the first four rounds and 2 for the final two rounds, for the other half the multiplier was 2 for the first four rounds and 1.2 for the final two; Nowell and Tinkler – 1.2 for all rounds; Seguino, Stevens, and Lutz – the multiplier was 2; Sell, Sell and Wilson, and Sell, Griffith, and Wilson – the multiplier was 3. Cadsby and Maynes varied the reward (5, 8, 10, or 15 tokens) for meeting the threshold.

[b]Pooling over all treatments. Brown-Kruse and Hummels had two treatments: Anonymity – subjects had no pre-experiment communication and did not know the identity of their fellow group members; Community – subjects had pre-experiment communication and did know the identity of their fellow group members. Sell and Wilson report three information-feedback treatments: no information, aggregate information, and individual information.

[c]In the Sell study, the total number of subjects numbered 254. Ten were excluded from the analysis. The Sell and Wilson study began with 96 subjects; four were excluded from the analysis. The Sell, Griffith, and Wilson study excluded 15 subjects from the original sample pool of 114.

[d]In the Sell and Sell, Griffith, and Wilson studies, subjects played in a group of four where the other three players were simulated.

[e]Hypothesis of like mean contributions rejected at $p \leqslant 0.05$.

[f]Regression analysis results indicate that contributions by all-female groups were significantly higher than contributions by all-male groups, $p = 0.10$.

[g]Regression analysis results indicate that contributions by female subjects were significantly higher than contributions by male subjects, $p = 0.01$.

The Brown-Kruse and Hummels (1993), Nowell and Tinkler (1994), and Sell studies provide information of average group contributions after each round of play.[6] Sell and Wilson (1991) varies the information conditions, with subjects receiving either no information, group information, or individual information. They find no significant interaction between the sex of the subject and the information treatment.

Finally, the studies vary in what the subjects know about their counterparts. In Sell and Wilson (1991), partners are anonymous; subjects do not know the sex of their counterpart. Sell (1997) and Sell, Griffith, and Wilson (1993) give the subjects only written stimuli (a list of their partners' first names) regarding their partners' sex. In the other three studies, stimuli are visual: all subjects are in the same room. Several of the studies use a computer interface, while others do not. Visual and written stimuli may not elicit the same response from subjects, and the computer interface may raise doubts in the subjects about the identity of their partner.

2. Ultimatum Experiments

Table 2 reports the findings from two ultimatum experiment studies. Solnick (2001) conducts a one-shot game under two treatments. In one treatment, subject anonymity is preserved; in the second, players are told the gender-revealing first name of their partner. Solnick's design employs the "strategy method": players simultaneously propose a division of the pie (if the player is the proposer) and the minimum acceptable division of the pie (if the player is the respondent).

Eckel and Grossman (2001) employ a repeated-play (eight periods) design. In each period subjects play with a different partner. Each subject plays each role (proposer and respondent) four times, though not in the same order. The sex of a subject's partner is made known by having a group of four proposers seated facing a group of four respondents. The design matches players with partners of their own gender (four men face four men, or four women face four women), partners of the opposite gender (four men face four women), or a mixed group (two women and two men face a similar group). Subjects are not told with whom they are matched. The Eckel and Grossman (2001) design employs the "game method": the proposer makes an offer which is presented to the respondent, who then decides whether to accept or reject.

There are considerable similarities in the Solnick (2001) and Eckel and Grossman (2001) results. Both studies find little difference in the overall mean offers made by men and women (46.7 percent of the pie for men versus 46.8 percent for women in the Solnick study and 36.5 versus 38.5 percent in the Eckel and Grossman, 2001, study). Both Solnick (2001) and Eckel and Grossman (2001) report that offers to women are,

---

[6] Sell manipulates the feedback received by subjects. Half of the subjects received feedback suggesting that other group members were playing cooperatively and half received feedback suggesting that other group members were playing non-cooperatively.

Table 2
Ultimatum experiments. Offers made by proposers, and rejection rates by respondents

Offers made by	Eckel and Grossman[a,b]		Solnick[a]	
	Mean offer (% of $5)	Rejection rate (%)	Mean offer (% of $10)	Rejection rate (%)
All subjects to all subjects	37.5	12.8	46.8	12.4
All subjects to men	38.2		48.9[c]	
All subjects to women	37.2		43.7[c]	
Men to all subjects	36.5	17.7	46.7	4.2
Men to men	36.6	18.8	47.3	4.5
Men to women	36.6	17.2	44.3	0.0
Women to all subjects	38.5	7.8	46.8	14.6
Women to men	39.8	9.4	51.3[d]	6.3
Women to women	37.8	3.1	43.1[d]	23.1

[a]Number of subjects: Eckel and Grossman – 96 subjects (each plays four rounds as proposer and four rounds as respondent; 384 proposer/respondent pairings); Solnick – 178 subjects (89 proposer/respondent pairings).
[b]Regression analysis results indicate that: (1) female respondents are significantly more likely to accept a given offer, *p*-value = 0.01; and (2) offers from female proposers are significantly less likely to be rejected than offers from male proposers, *p*-value = 0.01.
[c]Means test *p*-value = 0.08.
[d]Means test *p*-value = 0.08.

on average, lower than those made to men, regardless of the sex of the proposer (43.7 and 37.2 percent versus 48.9 and 38.2 percent, respectively).[7]

Where Solnick's (2001) and Eckel and Grossman's (2001) results differ dramatically is in the behavior of the respondents (see Table 2). While the overall rejection rates are similar (12.4 versus 12.8 percent, respectively), Solnick reports higher rejection rates of offers made by women, while Eckel and Grossman (2001) report higher rejection rates for offers made by men. Both Solnick's (2001) and Eckel and Grossman's (2001) results are significant. One of the most startling differences in the two results is the difference in rejection rates of offers made by women to women. In the Eckel and Grossman (2001) study, these offers were *least* likely to be rejected (3.1%), while in the Solnick (2001) study, these offers were the *most* likely to be rejected (23.1%).

There are two important differences between methodologies of Solnick (2001) and Eckel and Grossman (2001). One is the one-shot design versus repeated-play design. If subjects come to the experiment with no idea of what constitutes an "acceptable offer (minimum acceptable offer)," first-round results may reflect considerable "noise" as subjects experiment.

The second important difference is the risk differences and potential for being "exploited" faced by the respondent in the "strategy" versus "game method" design. In the

[7] The difference in the Solnick (2001) study is significant at the 8 percent level.

game method design, the respondent, knowing the proposer's offer, knows the outcome of the game once his decision is made. There is no risk and no potential for exploitation. In the strategy method design, the respondent faces the same risk as the proposer. Both must make a decision without knowing for certain the other's choice. The smaller is a proposer's offer and the higher is a respondent's minimum acceptable offer, the higher is the probability that both will receive nothing. There is both risk and potential for exploitation. Reactions of subjects to these risk differences may vary by sex, and by other characteristics of the subject pool.[8]

3. Dictator Experiments

The contradictory results of the prisoner's dilemma, public goods, and ultimatum experiments may be caused by failure to control for important environmental factors that might confound basic gender differences. The dictator experiment offers a simple design that removes possible confounding factors. The advantage of the dictator environment over others is that it eliminates considerations of strategic risk.

Eckel and Grossman (1996) test for differences in the behavior of men and women in what they called a "punishment" game, a variation on the dictator experiment. In the punishment game, subjects were given a choice of being paired with either, an "ungenerous" person (termed Type A in the experiment), an *unpaid* proposer from a previous experiment who had selected an $18/$2 split, or a "generous" person (Type B), an *unpaid* proposer from a previous experiment who had selected a $10/$10 split. The type of person that a subject chose to be paired with determined their own payoff. Half of the subjects randomly were assigned a "low relative price" payoff (sharing $12 with an ungenerous person or sharing $10 with a generous person), and half randomly were assigned a "high relative price" payoff (sharing $12 with an ungenerous person or sharing $8 with a generous person).

Eckel and Grossman (1996) report a significant difference in the behavior of men and women (see Table 3). Women were more responsive to changes in the parameters of the decision-making environment. In the low relative price treatment, women were about twice as likely as men to sacrifice $1 in earnings in order to choose to reward the generous person. In the high relative price treatment, however, women were slightly less likely than men to sacrifice $2 in earnings in order to choose the generous person. Men's choices did not vary with the price treatment.

Bolton and Katok (1995) and Eckel and Grossman (1998) employ the standard dictator experiment design, though their experimental environments differ. The choice set in one Bolton and Katok (1995) treatment was restricted to giving zero or 50 percent of the pie, and in all Bolton and Katok (1995) treatments the choice set was restricted to

---

[8] In particular, Eckel and Grossman (2001) report substantial differences in play by African–American subjects.

Table 3

Dictator punishment game. Subjects choose to split evenly a large amount with an "ungenerous" recipient or a small amount with a "generous" recipient[a]

Sex	Low relative payoff Percent choosing to		High relative payoff Percent choosing to	
	Split $10 with an ungenerous recipient (# of subjects)	Split $8 with a generous recipient (# of subjects)	Split $12 with an ungenerous recipient (# of subjects)	Split $8 with a generous recipient (# of subjects)
Men	60.7 (37)	39.3 (29)	59.2 (29)	40.8 (20)
Women	36.0 (18)	64.0 (32)	67.3 (35)	32.7 (17)
Total	49.5 (55)	50.5 (56)	63.4 (64)	36.6 (37)

Note. Subjects chose between splitting evenly a smaller amount of money with a recipient who had been a "generous" dictator in a previous game, and splitting evenly a larger amount of money with a recipient who had been an "ungenerous" dictator in a previous game.

[a]Three-way Goodman Placket test of interaction between Sex × Price × Punish: $\chi^2 = 5.64$, p-value $= 0.025$.

contributions of 50 percent of the pie or less. Eckel and Grossman (1998) only restrict the contributions of subjects to whole dollar increments.

Both Bolton and Katok (1995) and Eckel and Grossman (1998) find women donating more to their partners (see Table 4). However, Eckel and Grossman's (1998) difference is highly significant, while Bolton and Katok's (1995) is not. Bolton and Katok's (1995) restricted choice set may be the cause of the difference in results; it may unintentionally signal subjects that smaller donations are expected. If, as Bolton and Katok (1995) argue, '...when confronted with a choice of leaving more or less than they would freely choose, dictators choose less' (p. 290), then this restricted choice set may conceal any sex differences.

Andreoni and Vesterlund (2001) conduct a modified dictator experiment. Instead of making a decision over one choice set, a subject makes allocation decisions for eight different choice sets (a subject's payoff was determined randomly from the decisions made). Each choice set differed in the number of tokens to be divided and the value of a token to each subject.

Overall, women gave away more tokens than men (29.50 versus 25.74, respectively), but this varied considerably with the relative value of tokens to the two subjects (see Table 5; Andreoni and Vesterlund's, 2001, Table 2). Women's token donations varied little with the relative price of giving, while men's donations were more responsive. Increasing the value of a token to the recipient tended to increase men's level of donations. On average, partners of female subjects earned more than partners of male subjects ($2.60

Table 4
Dictator experiments. Women give more than men

	Mean donation as percent of total (# of subjects)	
	Bolton and Katok[a]	Eckel and Grossman[a]
Men	11.3%[b] (46)	8.2%[c] (60)
Women	12.3[b] (31)	16.0[c] (60)

[a]Bolton and Katok employ three treatments: (1) dictator plays with 10 different respondents, dividing $1 in $0.10 increments; (2) dictator plays with 10 different respondents, either taking the whole pie or taking half; and (3) dictator plays with 1 respondent, dividing $10 in $1 increments. In all three treatments, the dictator is restricted to giving away no more than one-half the pie. Eckel and Grossman employ one treatment: the dictator plays with one respondent, dividing $10 in $1 increments. There is no restriction on the amount the dictator may give away. Bolton and Katok did not impose anonymity between subjects and experimenters; Eckel and Grossman did.
[b]Bolton and Katok: means test p-value = 0.36.
[c]Eckel and Grossman: means test p-value < 0.01.

Table 5
Dictator games. Percent of endowment given to recipient under different budget constraints

	Token endowment	Relative price[a]	All subjects ($n = 116$)	Male subjects ($n = 84$)	Female subjects ($n = 32$)	Male–female difference
Budget 1	40	1/3	31.83	35.27	22.81	12.46
Budget 2	40	3	24.29	18.51	39.45	20.94
Budget 3	60	1/2	32.37	34.01	28.07	5.94
Budget 4	60	1	26.87	23.51	35.68	−12.17
Budget 5	60	2	25.10	20.08	38.28	−18.20
Budget 6	75	1/2	31.34	33.11	26.71	6.40
Budget 7	75	2	23.99	18.92	37.29	−18.37
Budget 8	100	1	25.15	22.35	32.50	−10.15
Average			27.62	25.72[b]	32.60[b]	−6.88

Source: Andreoni and Vesterlund (1997).
[a]Ratio of the value of a token to the decision maker relative to the value of a token to the recipient.
[b]Means test p-value = 0.75.

versus $2.56), and for five of the eight choice sets (budgets 1, 2, 4, 5, and 7) the earnings differential was significant. Only for budget 1 did the partners of male subjects earn significantly more than the partners of female subjects.

Andreoni and Vesterlund's (2001) outcomes for budgets 4 and 8 are directly comparable to the Eckel and Grossman (1998) and Bolton and Katok (1995) sessions. These two choice sets offer the same choice options as a simple dictator experiment. In both cases, the results are consistent with Eckel and Grossman's (1998) findings; women are significantly more generous than men, giving approximately 41% more than men for budget 4 and 26% more than men for budget 8.

4. Conclusions

While the results from public goods, ultimatum, and dictator experiments do not offer consistent evidence of behavioral differences between men and women, there are some intriguing patterns in the data. In those settings where subjects are exposed to risk – i.e., public goods experiments, as proposer in ultimatum experiments, and as respondent in ultimatum experiments employing the "strategy method" design – there is no significant evidence of systematic differences in the play of men and women. Results seem to depend on the details of the payoff structure and experimental procedure. In those settings where subjects are not exposed to risk – i.e., as respondent in ultimatum experiments employing the "game method" design and dictator games – systematic differences are revealed. The choices women make are less individually-oriented and more socially-oriented. This finding is conditioned by the level of risk, which points to an important avenue for future research. If men and women systematically differ in their responses to risk, then this has important implications for behavior.

References

Andreoni, J., Vesterlund, L. (2001). "Which is the fair sex? Gender differences in altruism". Quarterly Journal of Economics 116, 293–312.
Bolton, G.E., Katok, E. (1995). "An experimental test for gender differences in beneficent behavior". Economics Letters 48, 287–292.
Brown-Kruse, J., Hummels, D. (1993). "Gender effects in laboratory public goods contributions: Do individuals put their money where their mouth is?" Journal of Economic Behavior and Organization 22, 255–267.
Cadsby, C.B., Maynes, E. (1998). "Gender and free riding in a threshold public goods game: Experimental evidence". Journal of Economic Behavior and Organization 34, 603–620.
Eckel, C.C., Grossman, P.J. (1996). "The relative price of fairness: Gender differences in a punishment game". Journal of Economic Behavior and Organization 30, 143–158.
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.
Ingram, B.L., Berger, S.E. (1977). "Sex-role orientation, defensiveness, and competitiveness in women". Journal of Conflict Resolution 21, 501–518.
Nowell, C., Tinkler, S. (1994). "The influence of gender on the provision of a public good". Journal of Economic Behavior and Organization 25, 25–36.
Rapoport, A., Chammah, A.M. (1965). "Sex differences in factors contributing to the level of cooperation in the prisoner's dilemma game". Journal of Personality and Social Psychology 2, 831–838.

Seguino, S., Stevens, T., Lutz, M. (1996). "Gender and cooperative behaviour: Economic man rides alone". Feminist Economics 2, 1–21.

Sell, J. (1997). "Gender, strategies, and contributions to public goods". Social Psychology Quarterly 60, 252–265.

Sell, J., Wilson, R.K. (1991). "Levels of information and contributions to public goods". Social Forces 70, 107–124.

Sell, J., Griffith, W.I., Wilson, R.K. (1993). "Are women more cooperative than men in social dilemmas?" Social Psychology Quarterly 56, 211–222.

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