



## Of mice and men: Within gender variation in strategic behavior

Marco E. Castillo <sup>a,\*</sup>, Philip J. Cross <sup>b</sup>

<sup>a</sup> *School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332, USA*

<sup>b</sup> *Department of Economics, Georgetown University, Washington, DC 20057, USA*

Received 27 June 2006

Available online 4 March 2008

---

### Abstract

We study behavioral differences across and within genders in a family of ultimatum and dictator games. We find these differences are due not only to altruistic preferences but also beliefs about the strategic behavior of others. The behavior of men in strategic situations is not significantly more aggressive than women on average. But this average masks wide variation in intra-gender behavior. In particular, a sizable minority of males are “mice,” behaving timidly in strategic environments. Our experimental design shows that the standard ultimatum game can mask significant inter- and intra-gender differences in strategic behavior. These behavioral patterns in strategic environments are shown to be correlated with preferences for altruism in non-strategic settings. Such gender differences could well manifest themselves in real-world large-stakes transactions, such as salary negotiations.

© 2008 Elsevier Inc. All rights reserved.

*JEL classification:* C72; C78; C92; J16

*Keywords:* Noncooperative games; Experimental economics; Beliefs

---

---

\* Corresponding author.

*E-mail addresses:* [marco.castillo@pubpolicy.gatech.edu](mailto:marco.castillo@pubpolicy.gatech.edu) (M.E. Castillo), [pjc22@georgetown.edu](mailto:pjc22@georgetown.edu) (P.J. Cross).

## 1. Introduction

Economists hold a consensus view that a persistent gender gap in wages exists, even when conditioning on a wide range of human capital variables (Goldin, 1990). The primary focus in labor economics has been on higher average wages for males, but the variability of the wage distribution around the average is also greater for men. One source of these differences may be varying expectations and outcomes in strategic bargaining situations. There may exist wage-bargaining opportunities when employees are first hired as well as periodically throughout their careers. Wood et al. (1993) find males have higher means and also higher standard deviations in earnings one year out of law school, and also 15 years out. Gerhart and Rynes (1991) find similar patterns in the initial earnings of MBA graduates.

This paper presents experimental evidence of gender differences in bargaining similar to these findings in the labor-economics literature. We are able to identify these differences by observing behavior across different games. Importantly, collecting information on behavior across games is necessary for this purpose due to the fact that bargaining outcomes respond to both preferences and beliefs.

The emerging picture in the non-strategic literature is that systematic differences exist between men and women in terms of altruistic and risk preferences.<sup>1</sup> Men are less altruistic than women on average, but their demand for giving is more price-elastic than that of women (Andreoni and Vesterlund, 2001). Men are less risk averse than women, and men are less responsive to the strategic environment in which decisions are made (Croson and Gneezy, 2004). For instance, men outperform women when facing competition (Gneezy et al., 2003; Niederle and Vesterlund, 2006). One strategic setting where gender differences are not robust is in proposer behavior in simple games of proposal and response. Croson and Buchan (1999) report that in a trust game across a large variety of populations, there is no significant difference in the offers made by men and women. Castillo and Carter (2006) also failed to find a gender effect in the behavior of proposers in a trust game. Similarly, Solnick (2001) shows that the behavior of male and female proposers in an ultimatum game is not significantly different.<sup>2</sup> The lack of clear gender differences in simple proposal-response games is somewhat puzzling considering the findings that men are more selfish and less risk averse. One important issue is that data generated by an ultimatum game may confound preference-driven gender differences in strategic behavior with those that are belief-driven.

This paper disentangles preferences from beliefs by observing behavior in both ultimatum games and dictator games. If men are in general less altruistic and less risk averse than women, but we observe the behavior of men to be similar to women's on average when proposing in ultimatum games, then there is some fraction of men who are more pessimistic about the probability of acceptance of less generous offers. This paper presents across-game evidence consistent with the existence of such a group of male pessimists. These men are not more altruistic than women. They allocate less than women in dictator games but offer more than women in the ultimatum games.

We observe behavior in a family of ultimatum-game/dictator-game pairs, with endowments and relative prices of offers changing across pairs. The standard dictator and ultimatum games

---

<sup>1</sup> For risk preferences, see Jianakoplos and Bernasek (1998), Dwyer et al. (2002), Hudgens and Fatkin (1985), and Croson and Gneezy (2004). For altruistic preferences, see Eckel and Grossman (1998, 2001), and Croson and Gneezy (2004)

<sup>2</sup> An exception is Petrie (2004) who finds that men in a trust game pass larger amounts than women.

result when this relative price is set to 1. When the relative price exceeds 1, it is inexpensive for the proposer to give money to the responder. A family of dictator games of this type was analyzed in Andreoni and Miller's (2002) study of rational giving. We exploit the fact that an ultimatum game is simply a dictator game with the responder given the opportunity to reject the proposed allocation. Observing behavior in games with varying relative prices is necessary to identify any gender differences in substitutability between own and others' payoffs (Andreoni and Vesterlund, 2001).

The picture of men emerging from our experiments is nuanced. A large fraction of males behave more aggressively than females, as one would expect from previous studies. But there is also a sizable minority of males behaving timidly. The existence of both "mice" and "men" categories of male behavior leads to similar behavior across genders on average. Male behavior exhibits more variation than female behavior, both in strategic and non-strategic environments. This is not due solely to preferences for altruism, but also heterogeneity of beliefs in strategic behavior. Previous work eliciting beliefs in games have produced mixed results. Nyarko and Schotter (2002) show that beliefs are important in explaining behavior in games with unique Nash equilibria, but these beliefs are not those implied by bounded-rationality models. Costa-Gomes and Weizsacher (2005) provide evidence that players fail to play best-response to their elicited beliefs. And Croson (2000) shows that elicitation of beliefs significantly affects behavior in games, making players behave more strategically. In this paper we both infer beliefs from strategic choices and directly elicit beliefs through a post-experiment questionnaire. The directly elicited beliefs give some clues as to what is driving the wide diversity in male behavior.

The paper is organized as follows. The next section outlines the theoretical implications for behavior in ultimatum and dictator games with varying prices of giving. Section 3 describes our experiment. Section 4 presents the results, focusing on inter-gender and within-male variations in behavior. Section 5 addresses potential causes of these gender differences, and Section 6 concludes.

## 2. Theory

Let  $m$  be the endowment of dollars over which a proposer and responder are bargaining and let  $p$  be the relative price the proposer faces when passing money to the responder. We consider ultimatum games, where the proposer chooses an offer  $x \in \mathcal{X}$ , and the responder chooses  $y \in \mathcal{Y}^{\text{UG}} = \{0, 1\}$  indicating rejection or acceptance of the offer. The monetary payoffs for the proposer and responder are, respectively,

$$\pi_p(x, y) \equiv (m - px)y, \quad \text{and} \quad \pi_r(x, y) \equiv xy.$$

That is, the proposer's offer is (fully) accepted or (fully) rejected by the responder.

We suppose agents are expected utility maximizers; however, agents vary in their preferences and expectations. Standard game theory holds that an agent's utility is an increasing function of her own payoff and not a function of the payoffs of other agents. The standard theory yields a subgame perfect equilibrium where minimal offers are made by proposers and accepted by responders. However, researchers have long entertained non-standard behavior. In this case the proposer maximizes expected utility,

$$\max_{x \in \mathcal{X}} u(\pi_p(x, 1), \pi_r(x, 1)) \Pr(y = 1 | x) + u(\pi_p(x, 0), \pi_r(x, 0)) \Pr(y = 0 | x).$$

Note that  $\pi_p(x, 0) = \pi_r(x, 0) = 0$ . Suppose  $u(0, 0) = 0$ ,<sup>3</sup> and define  $Q(x) \equiv \Pr(y = 1 | x)$  as the subjective probability the proposer places on the responder accepting his offer, as a function of his choice.<sup>4</sup> The proposer's expected-utility maximization problem can now be written as

$$\max_{x \in \mathcal{X}} u(\pi_p(x, 1), \pi_r(x, 1)) Q(x).$$

We also implement a variation on the ultimatum game where the proposer's choice set remains  $\mathcal{X}$  but the responder's choice set is modified to  $\mathcal{Y}^{\text{DG}} = \{1\}$ . That is, the responder is forced to "choose" to accept the proposer's offer. This modification constitutes a dictator game. Thus in the dictator game the proposer simply chooses a division of payoffs along the budget line and the game ends. The above framework captures the dictator game by setting  $Q(x)$  identically equal to 1.

Should proposer behavior in a dictator game differ from her behavior in an ultimatum game with the same  $p$  and  $m$  parameters, then this difference is attributable to  $Q(x)$  not being constant at 1 in the ultimatum game. If  $Q(x)$  is non-decreasing in  $x$ , offers to responders in the ultimatum game will be no lower than in the corresponding dictator game. This holds regardless of the proposer's utility function,  $u(\pi_p(x, 1), \pi_r(x, 1))$ . A smaller offer to responders in the ultimatum game than in the corresponding dictator game implies the subjective probability function  $Q(x)$  is decreasing for some range of  $x$  values.<sup>5</sup> In contrast, if the proposer believes  $Q(x) = 1$  for all  $x$  in the ultimatum game, then her behavior will exactly coincide with that in the dictator game. This is the testable implication of the *fairness hypothesis* (Forsythe et al., 1994).

Absent assumptions about  $Q(x)$ , few restrictions on proposer behavior in ultimatum games are implied. Indeed, rationality alone has no testable hypotheses in games of proposal and response.<sup>6</sup> But placing some structure on  $Q(x)$  yields identifying power. One possibility with considerable appeal is to suppose each proposer's  $Q(x)$  is non-decreasing in  $x$ .

### 3. Experimental design

We implement versions of the two-player dictator and ultimatum games employed by Forsythe et al. (1994). The standard ultimatum game (UG) involves a *proposer* offering a division of  $m$  dollars between herself ( $\pi_p$ ) and the *responder* ( $\pi_r$ ), so that  $m = \pi_p + \pi_r$ . The responder then accepts or rejects the offered  $(\pi_p, \pi_r)$  allocation. If the responder accepts, his monetary payoff is  $\pi_r$  dollars and the proposer's monetary payoff is  $\pi_p$  dollars. If the responder rejects, both players receive a monetary payoff of zero dollars. The standard dictator game (DG) is an ultimatum game where the responder's choice set is restricted; he is no longer afforded the option of rejecting the proposal. Thus the proposer, or *dictator*, simply decides upon the division of the  $m$  dollars and the game ends.

In our experiment, subjects play four different UGs with varying endowments ( $m$ ) and relative prices of offers ( $p$ ). Subjects also play the four DGs with corresponding budgets. The subjects

<sup>3</sup> Gilboa and Schmeidler (2003) discuss axiomatic conditions under which preferences are unique up to a linear transformation, in which case we can assume without loss of generality that  $u(0, 0) = 0$ .

<sup>4</sup> The belief function  $Q(x)$  may also depend on parameters  $m$  and  $p$ . The notation leaves this implicit for the sake of clarity.

<sup>5</sup> For instance, if responders care about efficiency, it is possible that they reject offers that are favorable to them if this is wasteful. Our experiments include a game where this wasteful giving is possible.

<sup>6</sup> Gilboa and Schmeidler (2003) only imply that expected utility representations are linear in probabilities over responder's strategies.

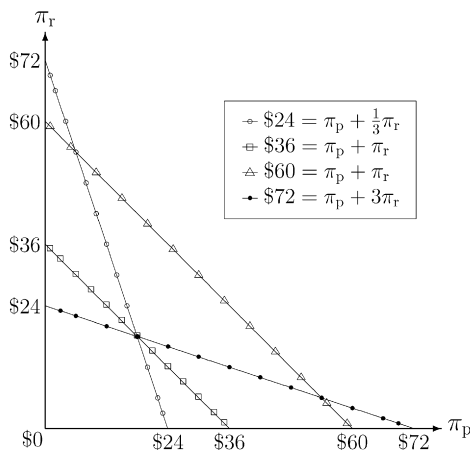


Fig. 1. Budget constraints faced by proposers.

are volunteers from undergraduate economics courses. Each subject makes choices assuming both the role of the proposer and that of the responder in the eight games. There is a fifty-fifty chance of ultimately being assigned the role of proposer or responder. Proposer choice sets are linear budget constraints,  $m = \pi_p + p\pi_r$ , discretized into 13 integer-valued dollar allocations. The four budget constraints are presented in Fig. 1.<sup>7</sup>

For example, for the UG with an endowment of  $m = \$24$  and a relative price of giving of  $p = 1/3$ , the strategy sets for the proposer and the responder are, respectively,

$$\mathcal{X} = \{3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 69\}, \quad \text{and} \quad \mathcal{Y}^{\text{UG}} = \{0, 1\}.$$

The monetary payoffs as a function of  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}^{\text{UG}}$  are  $\pi_p(x, y) = (24 - \frac{1}{3}x)y$  and  $\pi_r(x, y) = xy$ . The corresponding DG has the same choice set for the proposer,  $\mathcal{X}$ , but the responder’s choice set is restricted to  $\mathcal{Y}^{\text{DG}} = \{1\}$ . That is, the responder’s “choice” is always to accept the proposed offer.

The other three versions of the dictator and ultimatum games are defined likewise. For brevity we summarize these games by the convex, linear budget constraints listed in Table 1, not the actual discretized choice sets for proposers shown in Fig. 1. To make the choice sets more transparent, subjects were presented with the final dollar allocations rather than with budget constraints and endowments.<sup>8</sup> One hundred and twelve participants were recruited from undergraduate economics courses at Georgetown University. There were four experimental sessions with 29 participants each (Session 2 had 25 participants due to absenteeism). One participant in each session was chosen at random to be a monitor. The monitor made no decisions but rather verified to the other participants that the correct procedures were followed. Once the participants were assembled the instructions were read out loud, with participants reading along with their own copy. Subjects then filled out the experimental decision forms, and placed their completed decision forms in plain envelopes. In plain view these envelopes were collected, shuffled and

<sup>7</sup> Appendix A, on the *Games and Economic Behavior* website, gives the exact format under which proposers made their choices (see supplementary materials section).

<sup>8</sup> See Appendix A, on the *Games and Economic Behavior* website, for the decision sheets (see supplementary materials section).

randomly separated into two equal-sized piles, one for proposers and one for responders. Once proposer-responder pairs were formed, one of the 8 games was chosen at random and implemented. The forms were then taken to a nearby room to calculate payments. These payments, along with an \$8 attendance reimbursement, were placed in a private envelope with only the subject’s identification number on the outside. Another experimenter, not involved in the calculation of payments, handed out the envelopes to the participants, who were then escorted from the room. While payments were being calculated, subjects filled out a post-experiment questionnaire eliciting their understanding of the games, some expectations data, and some demographic covariates. The experiment lasted less than an hour, and participants earned an average of \$23.61 (s.e. \$2.11). Sixty-two of the subjects were male and 46 were female. One of the female subjects did not fill out her decision sheet for several of the games. The analysis excludes her, leaving 107 subjects.

### 4. Results

#### 4.1. Instrument check

The first panel of Table 1 shows average proposer behavior in the dictator and ultimatum games. By focusing on the two budgets where the relative price of giving is 1, we can compare

Table 1  
Average offers to responders,  $\pi_r$

<b>Overall:</b>	Dictator Game (DG)		Ultimatum Game (UG)			
$\$24 = \pi_p + \frac{1}{3}\pi_r$	\$8.66 (\$1.05)		\$15.70 (\$1.03)			
$\$36 = \pi_p + \pi_r$	\$6.14 (\$0.65)		\$12.14 (\$0.62)			
$\$60 = \pi_p + \pi_r$	\$8.49 (\$1.04)		\$18.89 (\$1.06)			
$\$72 = \pi_p + 3\pi_r$	\$4.69 (\$0.57)		\$10.73 (\$0.58)			
All four budgets	\$7.00 (\$0.43)		\$14.37 (\$0.45)			
<b>By gender:</b>	Males		Females			
	DG	UG	DG	UG		
$\$24 = \pi_p + \frac{1}{3}\pi_r$	\$8.08 (\$1.60)	\$16.65 (\$1.61)	\$9.47 (\$1.17)	\$14.40 (\$1.04)		
$\$36 = \pi_p + \pi_r$	\$4.47 (\$0.73)	\$12.03 (\$0.84)	\$8.45 (\$1.08)	\$12.29 (\$0.93)		
$\$60 = \pi_p + \pi_r$	\$6.03 (\$1.21)	\$17.85 (\$1.45)	\$11.87 (\$1.71)	\$20.31 (\$1.52)		
$\$72 = \pi_p + 3\pi_r$	\$3.18 (\$0.64)	\$10.69 (\$0.78)	\$6.78 (\$0.97)	\$10.80 (\$0.88)		
All four budgets	\$5.44 (\$0.56)	\$14.31 (\$0.64)	\$9.14 (\$0.64)	\$14.45 (\$0.62)		
<b>By gender and selfishness:</b>	Males			Females		
	Altruistic		Selfish	Altruistic		Selfish
	DG	UG	UG	DG	UG	UG
$\$24 = \pi_p + \frac{1}{3}\pi_r$	\$15.60 (\$3.48)	\$21.72 (\$2.68)	\$13.22 (\$1.81)	\$13.03 (\$1.44)	\$16.03 (\$1.33)	\$11.44 (\$1.43)
$\$36 = \pi_p + \pi_r$	\$9.60 (\$1.22)	\$15.24 (\$0.93)	\$9.86 (\$1.14)	\$12.55 (\$1.08)	\$13.10 (\$1.10)	\$10.81 (\$1.66)
$\$60 = \pi_p + \pi_r$	\$13.48 (\$2.31)	\$23.04 (\$1.94)	\$14.35 (\$1.86)	\$17.86 (\$1.88)	\$22.28 (\$1.78)	\$16.75 (\$2.65)
$\$72 = \pi_p + 3\pi_r$	\$6.40 (\$1.35)	\$13.92 (\$0.82)	\$8.51 (\$1.06)	\$9.97 (\$1.12)	\$11.45 (\$1.07)	\$9.63 (\$1.53)
All four budgets	\$11.27 (\$1.18)	\$18.48 (\$0.96)	\$11.49 (\$0.77)	\$13.35 (\$0.75)	\$15.72 (\$0.77)	\$12.16 (\$0.98)

Notes. 107 subjects, of which 62 are male (37 selfish and 25 altruistic) and 45 are female (16 selfish and 29 altruistic). Standard errors in parentheses.

our results to others in the literature. First consider the choices in DGs. When the price of giving is 1, proposers on average allocate to responders \$8.49 of a \$60 pie, and \$6.14 of a \$36 pie. So a unitary price sees proposers on average giving away 15.6 percent (s.e. 1.2%) of their endowments. This compares with 23 percent of budgets found to be given away by Forsythe et al. (1994), and 23 percent of budgets found to be given away by Andreoni and Miller (2002). Note that the budgets used in our experiment are larger than in these in previous studies,<sup>9</sup> and our findings of lower offers is consistent with our higher stakes.<sup>10</sup> Also, in contrast to these earlier studies, our experiment presented subjects with actual  $(\pi_p, \pi_r)$  pairs to choose from on their decision sheets.<sup>11</sup>

Now consider proposers' average offers in the four UGs. For the two unitary-price budgets, proposers offered \$12.14 of \$36 or 34 percent (s.e. 1.72%), and \$18.89 of \$60 or 31 percent (s.e. 1.77%) to responders. Compared to previous studies (Roth, 1995; Camerer, 2003), our subjects propose on average about the same in unitary-price UGs.

#### 4.2. Overall gender differences

Table 1 also shows the gender differences in the dictator and ultimatum games. There were 62 men and 45 women in total. As in Andreoni and Vesterlund (2001), men are less altruistic and less price-responsive than women. Table 1 shows men spend 12% of their budgets in the DG and this percentage remains constant as the price of giving varies. Women spend on average 21% of their budget in the DG and increase the share passed as the relative price of giving increases, so much so that the elasticity of substitution between  $\pi_p$  and  $\pi_r$  for women is less than one. Intriguingly, no gender differences in behavior are readily apparent in the UGs. On average, men and women pass roughly 33% of their budgets. While shares passed increase as the price of giving increases, there is no significant gender difference in strategic behavior.<sup>12</sup> This lack of gender difference in proposer-behavior in UGs mirrors the results of Croson and Gneezy (2004).

The second panel of Table 1 shows that the average allocations of males are significantly less generous than those of women in the DG ( $t = 4.32, p = 0.000$ ). But in the UG the average allocations do not significantly differ by gender ( $t = -0.16, p = 0.872$ ). This suggests a subset of males substantially change their behavior between the DG and the UG. The third panel of Table 1 separates subjects who allocate the minimum  $\pi_r$  in all DGs (termed *selfish*) from the remaining subjects (termed *altruistic*).

Twenty five of 62 males and 29 of 45 females are altruistic dictators, making males significantly less likely to be altruistic ( $t = -2.51, p = 0.013$ ). Among the selfish dictators, males do not give significantly more than females in the UG ( $t = -0.54, p = 0.593$ ). But a different pattern emerges among altruistic dictators. Altruistic males are less altruistic than altruistic females in the DG (allocating less on average). However, these same altruistic males are on average *more* generous with offers in the UG than altruistic females ( $t = 2.25, p = 0.025$ ). Thus

<sup>9</sup> Budgets in Forsythe et al. (1994) were worth \$10, budgets in Andreoni and Miller (2002) were between \$6 and \$10.

<sup>10</sup> Carpenter et al. (2005) find that the average allocations of dictators decreases from 33 percent to 25 percent when stakes are raised from \$10 to \$100. The sample in this study was small and this rise was not statistically significant.

<sup>11</sup> See Appendix A, on the *Games and Economic Behavior* website, for the decision sheets for the eight separate games (see supplementary materials section). Although these eight games appeared in the same order in all packets, subjects were free to go back and forth between the games to revise their decisions. We observed many subjects exercising this option of reviewing and modifying their decisions, and thus are confident order effects are not a cause for concern.

<sup>12</sup> All references to statistical significance in this section are at the 5 percent level.

altruistic males substantially change their behavior when moving from the DG to the UG. These DG-to-UG changes in behavior are examined in more detail below.

#### 4.3. Changes in proposer behavior between the dictator game and the ultimatum game

Table 2 examines in more detail subjects' changes in offers between the DGs and the corresponding UGs. Average DG-to-UG changes in offers to responders (denoted  $\Delta\pi_r$ ) are presented by gender, as well as statistics testing whether  $\Delta\pi_r$  is greater for males than for females on average. The first panel of Table 2 shows that on average males pass \$8.87 more in the UG than in the corresponding DG, compared with a \$5.31 increase for women, and that this gender difference is statistically significant. Also, considering behavior separately for each budget, males increase their offers more than females on average when facing each of the four budgets (although the  $p$ -value for the  $\$60 = \pi_p + \pi_r$  budget is 0.057, marginally exceeding 5 percent).

The second panel of Table 2 shows this gender difference broken down into altruistic and selfish categories. The average DG-to-UG increase in offers to responders is \$7.21 for the 25 altruistic males compared with \$2.36 for the 29 altruistic females. The greater DG-to-UG increase for males is statistically significant for each separate budget excepting  $\$24 = \pi_p + \frac{1}{3}\pi_r$  (where the  $p$ -value is 0.126). In contrast, gender differences in DG-to-UG increases in  $\pi_r$  are not statistically significant amongst selfish subjects.

#### 4.4. Intra-gender differences in behavior

Aside from gender-specific average DG-to-UG increases in offers, the variation around the average is also of interest. A simple measure of this intra-gender variation in behavior is the standard deviation. Table 3 gives estimates of the standard deviation of  $\Delta\pi_r$  by gender, and tests whether male behavior is more variable. The left-hand side of Table 3 shows this standard devi-

Table 2

Average DG-to-UG increase in offers ( $\Delta\pi_r$ ) for males and females (with  $t$ -statistic for test of average  $\Delta\pi_r$  being higher for males than for females)

Overall (62 males, 45 females)	Males		Females	$t$ -stat ( $p$ -value)		
$\$24 = \pi_p + \frac{1}{3}\pi_r$	\$8.56		\$4.93	2.04 (0.022)		
$\$36 = \pi_p + \pi_r$	\$7.56		\$3.84	2.66 (0.005)		
$\$60 = \pi_p + \pi_r$	\$11.82		\$8.44	1.60 (0.057)		
$\$72 = \pi_p + 3\pi_r$	\$7.52		\$4.02	2.76 (0.004)		
All four budgets	\$8.87		\$5.31	4.18 (0.000)		
<b>By selfishness:</b>	Altruistic (25 males, 29 females)			Selfish (37 males, 16 females)		
	Males	Females	$t$ -stat ( $p$ -value)	Males	Females	$t$ -stat ( $p$ -value)
$\$24 = \pi_p + \frac{1}{3}\pi_r$	\$6.12	\$3.00	1.17 (0.126)	\$10.22	\$8.44	0.77 (0.222)
$\$36 = \pi_p + \pi_r$	\$5.64	\$0.55	3.37 (0.001)	\$8.86	\$9.81	-0.47 (0.679)
$\$60 = \pi_p + \pi_r$	\$9.56	\$4.41	1.84 (0.037)	\$13.35	\$15.75	-0.74 (0.768)
$\$72 = \pi_p + 3\pi_r$	\$7.52	\$1.48	3.74 (0.000)	\$7.51	\$8.63	-0.60 (0.723)
All four budgets	\$7.21	\$2.36	4.34 (0.000)	\$9.99	\$10.66	-0.53 (0.702)



Table 3

Estimated standard deviations of  $\Delta\pi_r$  for males and females (with  $F$ -statistic for test of gender differences in standard deviations)

<b>All subjects</b> (62 males, 35 females)								
All four budgets			$\$24 = \pi_p + \frac{1}{3}\pi_r$			Remaining three budgets		
Males	Females	$F$ -stat	Males	Females	$F$ -stat	Males	Females	$F$ -stat
\$9.46 (\$0.40)	\$8.06 (\$0.37)	1.38 (0.011)	\$11.42 (\$0.88)	\$6.89 (\$0.62)	2.74 (0.000)	\$8.75 (\$0.41)	\$8.43 (\$0.47)	1.08 (0.325)
<b>Altruistic dictators</b> (25 males, 29 females)								
All four budgets			$\$24 = \pi_p + \frac{1}{3}\pi_r$			Remaining three budgets		
Males	Females	$F$ -stat	Males	Females	$F$ -stat	Males	Females	$F$ -stat
\$9.42 (\$0.59)	\$6.47 (\$0.37)	2.12 (0.000)	\$11.79 (\$1.35)	\$6.80 (\$0.75)	3.01 (0.003)	\$8.55 (\$0.62)	\$6.38 (\$0.44)	1.80 (0.005)
<b>Selfish dictators</b> (37 males, 16 females)								
All four budgets			$\$24 = \pi_p + \frac{1}{3}\pi_r$			Remaining three budgets		
Males	Females	$F$ -stat	Males	Females	$F$ -stat	Males	Females	$F$ -stat
\$9.36 (\$0.50)	\$7.94 (\$0.61)	1.39 (0.070)	\$11.01 (\$1.10)	\$5.72 (\$0.81)	3.71 (0.004)	\$8.80 (\$0.53)	\$8.47 (\$0.74)	1.08 (0.395)

Notes. Standard errors in parentheses below estimates.  $p$ -values in parentheses following  $F$ -statistics.

ation for males and females over all four budgets. Altruistic males exhibit significantly greater variation than altruistic females.<sup>13</sup>

So while selfish males and females exhibit no significant differences in  $\Delta\pi_r$  on average (as shown in Table 2), variation around this average is significantly greater for males (at least at the 10-percent level). And compared with altruistic females, altruistic males have a higher  $\Delta\pi_r$  on average (see Table 2) and a greater dispersion around this average.

We can examine the effect of the economic parameters  $p$  and  $m$  on this intra-gender variation in behavior by isolating each budget constraint. The middle and right-hand sides of Table 3 show the gender differences in the standard deviation of  $\Delta\pi_r$  for the  $\$24 = \pi_p + \frac{1}{3}\pi_r$  budget separate from the other three budgets. This budget has the smallest relative price and the smallest endowment in the experiment. For the altruists the gender difference in the standard deviation is apparent both at the low ( $p, m$ ) budget and the higher ( $p, m$ ) budgets. For the selfish, the gender difference is significant at the low ( $p, m$ ) budget but insignificant at the higher ( $p, m$ ) budgets.<sup>14</sup> Thus, while Table 2 shows no gender difference in  $\Delta\pi_r$  on average for any budget, Table 3 shows male behavior is significantly more variable at the low ( $p, m$ ) budget, but not more variable at the higher ( $p, m$ ) budgets. A low price of giving and a low endowment elicits a wider range of behavior among males relative to females, both for the altruistic and for the selfish.

<sup>13</sup> For selfish dictators the gender difference in the standard deviation of  $\Delta\pi_r$  is statistically insignificant, although the  $p$ -value of 0.070 is close to 5 percent.

<sup>14</sup> This statistical insignificance of the gender difference continues to hold when these three higher ( $p, m$ ) budgets are examined separately ( $p$ -values of 0.441, 0.413 and 0.413, respectively).

## 5. What drives these gender differences?

Having identified gender differences both in average behavior and in the variability of behavior, it seems natural to ask what is driving these differences. Are the results reported in Section 4 due to men and women differing in some systematic, observable way? Our experimental data, and additional covariate data beyond gender that we collected, allow (partial) investigation of this question. Addressed below are three potentially important factors that have attracted attention in the literature.

### 5.1. *The fairness hypothesis*

Suppose subjects are capable of pure altruism but otherwise exhibit standard game-theoretic behavior—in particular, responders accept all offers with probability 1 and proposers realize this. A testable implication of this fairness hypothesis is that UG-to-DG changes in proposals are zero. 12 of 62 males and 7 of 45 females behave according to the fairness hypothesis, an insignificant gender difference ( $t = 0.50$ ,  $p = 0.616$ ). Focusing on subjects that do not conform to the fairness hypothesis leaves the results in Section 4 essentially unchanged.

Among those not conforming to the fairness hypothesis, DG-to-UG increases in  $\pi_r$  average \$11.00 for males versus \$6.29 for females, a significant difference ( $t = 4.95$ ,  $p = 0.000$ ). For altruistic dictators not conforming to the fairness hypothesis (22 males and 24 females), this significant gender difference in average  $\Delta\pi_r$  persists ( $t = 4.26$ ,  $p = 0.000$ ). But for selfish dictators not conforming to the fairness hypothesis (29 males and 14 females) there is no significant gender difference on average ( $t = 0.80$ ,  $p = 0.211$ ). These findings mirror those in Table 2 above (which did not condition on the fairness hypothesis). The fairness hypothesis does not provide a systematic way to account for the gender difference found in Section 4.

### 5.2. *Agent behavior in the responder role*

Gender differences in agents' behavior in the proposer role might be driven by (or, at least, correlated with) the way these same agents behave in the responder role. We collected experimental data on behavior in the responder role, allowing the proportions of males and females who accept even the least-generous offers to be compared. Interestingly, altruistic males are not significantly greater accepters of these least-generous offers ( $t = -0.29$ ,  $p = 0.615$ ) than altruistic females on average. But a gender difference is present among the selfish dictators ( $t = 2.12$ ,  $p = 0.018$ ). This is the opposite of what we would expect if agents' responder-behavior were, in fact, driving average gender difference among proposers.

### 5.3. *Directly elicited beliefs*

Subjects' beliefs about the probability of rejection for several different proposals were elicited in a post-experiment questionnaire.<sup>15</sup> The questionnaire yielded data indicating whether, for 12 separate ( $\pi_p, \pi_r$ ) offers that were feasible choices in the experiment, the proposer believed the

<sup>15</sup> These elicited beliefs are valid if the responses are not simply a justification of past behavior. These responses are informative when they provide counterfactual information about offers rarely proposed in the experiment. To prevent bias, we did not elicit subjects' beliefs until all their decisions had been made.

probability the offer would be rejected was greater than 70 percent. We linearly fitted this indicator of pessimistic beliefs to  $\frac{\pi_r}{\pi_p + \pi_r}$  for both genders. This fit captures potential gender differences in the sensitivity of beliefs to the relative generosity of offers. Among altruistic dictators, males believe rejection is significantly more sensitive to the relative stinginess of offers than do females ( $t = 2.44$ ,  $p = 0.007$ ). Among the selfish there is no significant gender difference in the sensitivity of elicited beliefs ( $t = -1.41$ ,  $p = 0.921$ ). Notice that these gender differences in directly elicited beliefs mirror the Table 2 findings on DG-to-UG changes in offers.

Thus we have a partial explanation of what makes a sizable minority of males behave as “mice.” These “mice” are altruistic dictators who fear that even modestly stingy offers in the UG will be rejected with high probability. And another group of males constitute the “men,” behaving selfishly when dictators and not fearing rejection when proposers in the UG.

## 6. Conclusion

We identify gender differences in behavior by exploiting a novel experimental design. Observing choice in dictator and ultimatum games with varying relative prices reveals the strategic behavior of males exhibits wide variability, with important implications for real-world decision making. This behavior is consistent with the fact that the distribution of wages for males is more dispersed than that for females. And, more broadly, it is consistent with the very widely dispersed achievement levels for men—they are over-represented both in boardrooms as CEOs and behind bars as incarcerated criminals. We show these behavioral differences are partially explained by differences in expectations.

## Acknowledgments

We thank Georgetown University for financial support. And we would like to thank, without implicating, Alison Cantatore, Jeff Dominitz, Helga Fehr-Duda, Alex James, Deniz Kurtel, Matt Maddox, Ragan Petrie, Helena Poon two reviewers and a co-editor.

## Supplementary Appendix

Supplementary material associated with this article can be found, in the online version, at DOI: 10.1016/j.geb.2008.01.009.

## References

- Andreoni, J., Miller, J., 2002. Giving according to GARP: An experimental test of the consistency of preference for altruism. *Econometrica* 70 (3), 737–753.
- Andreoni, J., Vesterlund, L., 2001. Which is the fair sex? Gender differences in altruism. *Quart. J. Econ.* 116, 293–312.
- Camerer, C., 2003. Social preferences in dictator, ultimatum, and trust games. In: *Behavioral Game Theory*. Princeton Univ. Press, Princeton, NJ.
- Carpenter, J., Verhoogen, E., Burks, S., 2005. The effect of stakes in distribution experiments. *Econ. Letters* 86, 393–398.
- Castillo, M., Carter, M., 2006. Identifying social effects with economic field experiments. Mimeo, Georgia Institute of Technology.
- Costa-Gomes, M., Weizsacher, G., 2005. Stated beliefs and play in normal-form games. Mimeo, London School of Economics.
- Croson, R., 2000. Thinking like a game theorist: Factors affecting the frequency of equilibrium play. *J. Econ. Behav. Organ.* 41, 299–314.

- Croson, R., Buchan, N., 1999. Gender and culture: International experimental evidence from trust games. *Amer. Econ. Rev.* 89, 386–391.
- Croson, R., Gneezy, U., 2004. Gender differences in preferences. Mimeo, University of Chicago.
- Dwyer, P., Gilkeson, J., List, J., 2002. Gender differences in revealed risk taking: Evidence from mutual fund investors. *Econ. Letters* 76, 151–158.
- Eckel, C., Grossman, P., 1998. Are women less selfish than men? Evidence from dictator experiments. *Econ. J.* 108, 726–735.
- Eckel, C., Grossman, P., 2001. Chivalry and solidarity in ultimatum games. *Econ. Inquiry* 39, 171–188.
- Forsythe, R., Horowitz, J., Savin, N., Sefton, M., 1994. Fairness in simple bargaining games. *Games Econ. Behav.* 6, 347–369.
- Gerhart, B., Rynes, S., 1991. Determinants and consequences of salary negotiations by male and female MBA graduates. *J. Appl. Psychol.* 76 (2), 256–262.
- Gilboa, I., Schmeidler, D., 2003. A derivation of expected utility maximization in the context of a game. *Games Econ. Behav.* 44, 184–194.
- Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: Gender differences. *Quart. J. Econ.* 118 (3), 1049–1074.
- Goldin, C., 1990. *Understanding the Gender Gap*. Oxford Univ. Press, New York.
- Hudgens, G., Fatkin, L., 1985. Sex differences in risk taking: Repeated sessions on a computer-simulated task. *J. Psychol.* 119, 197–206.
- Jianakoplos, N., Bernasek, A., 1998. Are women more risk averse? *Econ. Inquiry* 34, 620–630.
- Niederle, M., Vesterlund, L., 2006. Do women shy away from competition? Mimeo, University of Pittsburgh.
- Nyarko, Y., Schotter, A., 2002. An experimental study of belief learning using elicited beliefs. *Econometrica* 70, 971–1005.
- Petrie, R., 2004. Trusting appearances and reciprocating looks: Experimental evidence on gender and race preferences. Mimeo, Georgia State University.
- Roth, A., 1995. Bargaining experiments. In: Kagel, J., Roth, A. (Eds.), *Handbook of Experimental Economics*. Princeton Univ. Press, Princeton, NJ.
- Solnick, S., 2001. Gender differences in the ultimatum game. *Econ. Inquiry* 39, 189–200.
- Wood, R., Corcoran, M., Courant, P., 1993. Pay differences among the highly paid: The male–female earnings gap in lawyers' salaries. *J. Lab. Econ.* 11 (3), 417–441.