Tests of Difference Aversion to Explain Anomalies in Simple Bargaining Games[1]

John H. Kagel
Department of Economics
Ohio State University

and

Katherine Wolfe
Department of Economics
University of Pittsburgh

7/19/00

Abstract

We investigate two recent models incorporating fairness considerations into the economics literature based on agents' concerns about the distribution of payoffs between themselves and others, along with their own absolute earnings (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). We extend the models' predictions to a three- person ultimatum game in which one person allocates a sum of money to two others, one of which is randomly chosen to accept or reject the offer. Rejection gives both the responder and the proposer zero income, but yields a consolation prize for the non-responder. If distributional concerns are the primary driving force behind rejections of offers in ultimatum games, as both models argue, rejections should be eliminated, or sharply reduced with a positive income consolation prize, since rejection establishes strong income inequality between the responder and the non-responder, with no income to offset the resulting inequality. The data show essentially no reductions in rejection rates, holding offers constant, with and without positive consolation prizes. We briefly explore alternative explanations for the results reported.

[1] This research has been partially supported by a grant from the Economics and DRMS Divisions of the National Science Foundation. We thank Bob Slonim and David Cooper for many thoughtful discussions on the issues under study here and suggestions regarding the experimental design. We thank seminar participants at SUNY Stony Brook, the Economic Science Association Meetings at Lake Tahoe and the ASSA meetings in Boston for helpful comments. Any errors are our own responsibility.

One of the more exciting recent developments in the economics literature has been the effort to incorporate notions of fairness into economic models. Much of this work has been motivated by results from ultimatum game and dictator game experiments which clearly show that economic agents base their decisions on more than their own narrow self-interest. Two of the most ambitious attempts along these lines have been papers by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000). Both models are based on the notion that in addition to the usual concerns about own payoffs, agents are motivated to reduce differences in payoffs between themselves and others. Further, these distributional concerns are largely self-centered, with greater distaste for having lower, rather than higher, earnings than others. What is most impressive about these two closely related models is their ability to organize a very large body of experimental data, both for games where distributional/fairness considerations appear to play a very important role (such as dictator and ultimatum games) and for games in which the standard, narrow self-interest model organizes the data quite well (for example, market games with proposer competition or public good games).

As with any model designed to fit an existing set of data, these two models have embedded within them new and interesting predictive consequences that have yet to be investigated. As such the question becomes how well do these models' do in these new situations and what sorts of modifications, if any, are needed to account for the additional data? The present paper is designed to test the Fehr-Schmidt (FS) and Bolton and Ockenfels (BO) models through a very simple extension of the standard ultimatum game to three players. The test is central to both models because it is (1) based on the same platform (the ultimatum game) central to the models' development and (2) focuses directly on the models' explanation for rejections of offers in the ultimatum game, aversion to income inequality.

Our experiment works as follows: Player X offers to split a sum of money between herself and players Y and Z. One of the latter is chosen at random to accept or reject the offer. If the responder accepts, than the proposed allocation is binding, as in the standard ultimatum game. However, if the responder rejects, both she and X get zero payoffs, but the non-responder receives a payoff. In the case of a sufficiently large positive payoff for the non-responder, both models call for *all* positive offers to be accepted as rejection calls for giving up own income while simultaneously establishing strong income inequality between the responder and the third player.

We reject both strong and weak versions of the FS and BO models. There are frequent rejections of offers when both models call for acceptance. In addition, the effect of positive payoffs for the non-responder on the probability of responders accepting offers is small, and fails to increase monotonically with size of the "consolation prize" as a weak version of both models would suggest. Further, the introduction of negative payoffs for the non-responder in the case of rejections has no effect as well. Modifications of FS and BO models required to account for our data consist of either (1) specifying a very narrow reference group against which income inequality is compared, which sharply limits their predictive consequences and/or (2) attaching a much greater role to negative reciprocity (Rabin, 1993; Levine, 1998) than either model does.

Observed Anomalies in Dictator and Ultimatum games

In an ultimatum game one bargainer (player X) makes a proposal on how to divide a sum of money with another player (player Y) who has the opportunity to either accept or reject the proposed division. If the proposed division is accepted, each player earns the amount proposed. If the proposal is rejected, then both players earn zero. The unique sub-game perfect equilibrium outcome for this game (under the assumption that players only seek to maximize their own income) is for player X to allocate all of the money to herself (or almost all of it, if payoffs are discrete) and for player Y to accept this proposed split. However, experimental evidence clearly shows that in games with symmetric payoffs and full information about both players' payoffs, that offers of positive amounts are routinely rejected. For example, Roth *et al.* (1991) found that positive offers were rejected 28% of the time in their U.S. $10 ultimatum experiment, with relatively low offers rejected quite frequently (in the U.S., offers of 30% or less were rejected at least half the time).

Dictator games differ from the ultimatum game by eliminating the opportunity for player Y to reject X's offer. Rather, player X dictates her offer to player Y who has no choice but to accept the proposed offer. In games of this sort, dictators commonly offer more than the smallest possible amount. For example, Forsythe *et al.* (1994) found that more than half of their subjects gave positive amounts to player Y when splitting $5 or $10. (However, the amounts allocated to player Y are considerably less, on average, than those offered in corresponding ultimatum games.)

Results from ultimatum and dictator game experiments clearly violate the standard (own) self-interest model of economics, as this model predicts the sub-game perfect equilibrium outcome in the ultimatum game and zero offers in the dictator game. Results from these games lie at the heart of the FS and BO models.

Using Distributional Considerations to Explain Anomalies in Two-Person Dictator and Ultimatum games

Although the FS and BO models differ in a number of important details, central to both is the notion that at least some agents in the population are motivated by both own income earnings and concerns about the distribution of income between themselves and other players in the game. This concern for the distribution of income expresses itself through these agents attaching some weight in their utility function to the distribution of earnings, with particular emphasis placed on *own* relative payoffs. Details, as they relate to our experiment, follow.

Bolton and Ockenfels (1998, 2000) proposed a utility function which measures fairness as the individual's share of the total payoff:

$$U_i(\mathbf{x}) = \mathrm{v}_i(\mathrm{x}_i, s_i)$$

$$\text{where} \quad s_i(\mathrm{c}, \mathrm{x}_i) = \begin{cases} \dfrac{\mathrm{x}_i}{\mathrm{c}}, & \text{if } \mathrm{c} > 0 \\ 1/n, & \text{if } \mathrm{c} = 0 \end{cases}$$

x_i is player I s payoff
$\mathbf{x} = (x_1, x_2, ...x_n)$ is the payoff vector
c is the sum of all players payoffs
ó$_i$ is player i s relative payoff share, ó$_i$ 0 [0,1]

v_{i1} \$ 0
v_{i11} # 0
$v_{i2} = 0$ for ó$_i$ = 1/n
$v_{i22} < 0$

The BO model readily accounts for the outcomes of two-person dictator games and ultimatum games. In a two-person dictator game the dictator, X, will choose an allocation (x*, y*) between himself and player Y which maximizes U_x (x, y). BO assume that the optimal share of the total payoff for player i is $ó_i$ (x*) ≥ 1/n, so that the dictator will never give away more than half the total amount to be allocated (which, indeed, rarely happens). In a two-person ultimatum game where X proposes and Y responds, Y will have a rejection threshold r_y, where, for y/c < r_y, U_y (0,0) = v_y (0,1/2) > U_y (x, y) = v_y (y, y/c); i.e., Y gets more utility from having an even share of nothing than from having a positive payoff but a very small share of the pie.

Fehr and Schmidt (1999) proposed a utility function which measures fairness by comparing the individual's payoff to other's payoffs:

.

$$U_i(\mathbf{x}) = x_i - a_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - b_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_i - x_j, 0\}$$

$$\text{where} \quad b_i \leq a_i,\ 0 \leq b_i \leq 1$$

They separate, disadvantageous inequality where the individual receives a lower payoff than some others do from advantageous inequality where the individual receives a higher payoff. They use an additively separable utility function where both types of inequality are negatively weighted, with disadvantageous inequality more heavily weighted than advantageous inequality. When everyone is receiving the same amount, there is no disutility from inequality.

The FS model also accounts for the outcomes of two-person ultimatum games. In a two-person ultimatum game where X proposes and Y responds, Y will have a rejection threshold r_y, where, for y/c < r_y, U_y (0,0) = 0 > U_y (x, y) = y - $á_y$ (x-y); i.e., the disutility of the inequality [$á_y$ (x-y)] outweighs the value of the positive payoff y. FS modify their utility function by allowing the disutility components to be nonlinear in x_i -x_j to account for dictator game results. In the dictator game, the dictator X chooses the allocation (x*, y*) which maximizes U_x (x, y). Note that X will never choose y > x because for any point where y > x, by increasing x, X can increase utility from his or her direct payoff and decrease the disutility from the disadvantageous inequality. So, as in BO, the dictator will never give herself less than half of the pie.

Three-person ultimatum games – some theoretical predictions

For two-person ultimatum games both BO and FS are able to organize the basic experimental outcomes. But a three-person game allows for more possibilities in varying payoffs and inequality, which provides the basis for our experimental design. Güth and van Damme (1998; hereafter GvD) first suggested the possibility of using a three-person ultimatum game to examine fairness issues in ultimatum games.

In GvD player X proposed a split (x, y, z) of 120 points for players X, Y, and Z. Player Y then accepts or rejects this split. If player Y accepts, the money is allocated as player X proposed. If player Y rejects, all three players receive zero. GvD found that X took advantage of Z's "dummy" status, essentially dividing the money between herself and Y, with Z receiving very low offers. Further, when Y had full information about the proposed split, these offers were virtually never rejected.

Bolton and Ockenfels (1998) argue that these results provide strong support for their model of self-centered inequality aversion. Consistent with their model's prediction, offers of 1/3 or more, the social reference norm applicable to this game, were never rejected. In contrast, offers of less than 40% of the pie are frequently rejected in two-person ultimatum games. In addition, they argue that concern for Z's payoff had no impact on the rejection decision, since proposers allocated such small amounts to the "dummy" and "there is not a single rejection that can be attributed to a low share for the dummy (GvD, p. 230)." These low rejection rates can be rationalized within the FS model by a utility function where the weight on advantageous inequality (â) is not large enough for Y to reject X's offers on account of the small amounts offered to Z, in conjunction with the relatively large payoffs offered to Y.

In what follows we modify the GvD game to provide a much stronger and more demanding test of the FS and BO models.

Experiment 1: A modified three person ultimatum game

We make two basic changes to the GvD game: First, the responding player is randomly selected to be either Y or Z in each round, with the actual selection determined *after* X has made her allocation. Allocating between Y and Z from behind a veil of ignorance as to which one will be responding prevents X from marginalizing the non-responding player in order to buy off the responding player. Other things equal this should result in lower offers to the responding player than in GvD, and the lower the allocation the higher the probability of rejection in a standard ultimatum game. Second, the rejection outcome for the non-responding player is varied so that she does not always get a zero payoff in cases where the responding player rejects X's allocation. The payoff for the non-responding player varies between experimental sessions taking on values of $0, $1, $3 and $12, all in games where X allocates $15. Changing the rejection outcome so that the non-responding player receives a positive "consolation prize" eliminates all possibility of rejections of positive offers within the BO model and, for the $12 consolation prize, should effectively eliminate all rejections within the FS model. We elaborate on these differential predictions for the FS and BO models. In doing so we will refer to the responding player as Y and to the non-responding player as Z, regardless of whether or not Y or Z was actually chosen to be the responding player.

For BO the existence of a positive consolation prize should eliminate *all* rejections, no matter how small the payoff to Y. For example, consider a proposed allocation of (9, 3, 3). If Y accepts the allocation she obtains utility v_y (3, 1/5). However, if she rejects, and the non-responder (Z) receives a positive consolation prize, no matter how small, then Y's utility would be reduced to v_y (0,0), so that accepting clearly *dominates* rejecting. In contrast, without the positive consolation prize, Y's utility would be v_y (0, 1/3), which permits rejection since it may be greater than v_y (3, 1/5).

In contrast to BO, the FS model permits some rejections with a positive consolation prize. For example, consider the proposed allocation (9, 3, 3), so that there is no advantageous inequality between Y and Z, but disadvantageous inequality between X and Y. With any consolation prize less than \$6 the amount of disadvantageous inequality that would result from a rejection of X's offer is less than the disadvantageous inequality resulting from accepting the offer. As such the offer can be accepted or rejected without violating the model's predictions, depending on the weight that Y places on the disadvantageous inequality that will result from accepting the offer. For example, with a \$3 consolation prize, acceptance yields utility of $3 – 0.5\alpha_y(6)$, while rejection yields $0 – 0.5\alpha_y(3)$ so that depending on the value of α, Y will accept or reject the offer. However, with this same allocation and a consolation prize for Z greater than \$6 (6+), the FS model no longer permits rejections, since Y forgoes the benefit of a positive payoff, and the disadvantageous inequality resulting from rejection is now greater than had she accepted the offer; i.e., $3 – 0.5\alpha_y(6)$ *dominates* $0 – 0.5\alpha_y(6+)$. Further, once the amount of disadvantageous inequality from rejection is greater than from acceptance the offer must be accepted, regardless of the amount of advantageous inequality the offer provides Y compared to Z. [2] The $12 consolation prize is designed, in part, to be large enough that virtually all allocations will have to be accepted according to the FS model.

Experimental Procedures

Player X divided $15 among three people. There were four sessions, differing only in the consolation prize that the non-responding player got if X's offer was rejected: $12, $3, $1, $0. The $1 session was run first, with a slight modification in the procedures as noted below.

For each session, 30 subjects (27 for the $3 session) were randomly assigned to the roles of X, Y, and Z, with these roles fixed throughout the session. Each session had 10 rounds with the responders X was paired with rotating after each round and with the pairings of Y and Z rotating as well. In the session with 27 subjects there was one rematch, but subject ID numbers changed between rounds so that there was no way subjects could identify that they were paired with the same partners as in a previous round, and the rematch was selected at random. X was required to propose a minimum of $0.10 to both Y and Z. The minimum ensured that in choosing to reject, the responding player is actually giving up some positive payoff.

In the $1 session, the responding player was selected prior to Y or Z seeing the allocations and then only the selected players got to see the offer and respond to it. In the other sessions, both Y and Z players received a copy of the proposed allocation, made their decisions, and then one set of decisions

[2] For example, consider the following extreme alternative to the example in the text, (9, 6, 0). Acceptance yields utility of $6 - 0.5\alpha_y(3) - 0.5\beta_y(6)$ and rejection with a consolation prize greater than 3 (3+) yields $0 - .5\beta(3+)$. But under FS's assumption that $0 \leq \beta \leq 1$ the negative effect of the advantageous inequality associated with acceptance cannot offset the loss of income ($6) associated with rejection. As we will be shown, relaxing this assumption has very little impact on our analysis since the overwhelming majority of offers provided Y and Z with the same amount of money.

(Ys or Zs) was randomly selected to determine the outcome of that round. However, only the selected decisions were returned to X players. Further, Y and Z players only knew what decision they made, but not the other's choice. So, for example, if Ys' decisions were chosen as binding in a given round, Xs only got to see Y's decision for the Y player they were paired with for that round, and Zs would not know their allocation for that round until the experiment ended (if that round was chosen for payment). These procedures preserve the traditional information structure underlying two-person ultimatum games whereby subjects only know their own outcomes. Finally, by having both Y and Z decide prior to determining which choice is binding provides twice as many responder observations.

To ensure that subjects were familiar with the rules of the game, especially the consequences of rejecting an offer, there was a pre-test in which all players were required to evaluate the consequences of an offer being accepted or rejected, including the consolation prize that Z would get in case of rejection.[3] Subjects were asked to record the outcome of each round. Payoffs were made at the end of the session for one randomly selected round. Subjects also received a show up fee of $5. Sessions were conducted by hand, with the typical session lasting a little over one hour. Subjects were recruited from a broad cross-section of undergraduate and graduate students (along with an occasional staff member) at the University of Pittsburgh and Carnegie Mellon University.

Results from Experiment 1:

Responder Behavior:

Table 1 and Figure 1 summarize the offers received by the responding players and their decisions. The rejection rate for the $0 consolation prize session is 21% compared to an overall rejection rate of 19%.

[3] The X player's quiz had them fill in an offer and then evaluate the consequences of acceptance or rejection. The Y

In all treatments the rejections follow the same pattern as in the two-person ultimatum game, being concentrated on those offers that are less than half of the social norm of an even split among all players (1/6 of the pie; less than \$2.50): Eighty percent of all offers less than \$2.50 were rejected, compared to 12% of all offers greater than or equal to \$2.50.

[Insert Table 1 and Fig 1 here]

We can immediately reject the point predictions of both the BO and FS models. Contrary to the BO prediction of zero rejections, there are non-negligible rejection rates for all positive consolation prize treatments, with these rejection rates approximately the same as with the \$0 consolation prize treatment. For the FS model, the cleanest test is with the \$12 consolation prize, since this treatment provides the largest number of observations for which the disadvantageous inequality associated with a rejection is greater than the disadvantageous inequality associated with X's offer (97% of all offers). Of these, 20% (39/193) were rejected, essentially the same overall rejection rate (21%) as in the zero consolation prize treatment.[4] Further, the rejection rates fall right in line with the overall rejection rates in the other treatments (75% for offers between \$1.75-\$2.24, 28% for offers between \$2.75-\$3.24, 10% for offers between \$3.75-\$4.24, and 0% for offers greater than or equal to \$4.25; see Table 1).[5]

The point predictions for these two models are indeed tough criteria to satisfy. Further, both sets of authors note that there is some role for other factors such as negative reciprocity (a willingness to

---

and Z player's quiz gave (7, 5, 3) as a sample offer with them evaluating the consequences of acceptance or rejection.

[4] Although last period data are quite thin, they are right in line with these results as well: 67% (2/3) last period offers of less than \$2.50 were rejected under the zero consolation prize treatment versus 100% (2/2) with positive consolation prizes. For offers greater than or equal to \$2.50 6% (1/17) were rejected with no consolation prize versus 11% (5/46) with positive consolation prizes.

[5] The other treatments involve much weaker tests of the FS model. For example, with the \$3 consolation prize, 38.3% of all offers had disadvantageous inequality of \$3 or less, with very few of these offers rejected (2.9%). However, these offers give the responder at least \$4, which is very close to the social norm of an equal split, so that there is very little likelihood of rejection under any circumstances.

punish others who knowingly harm you) to affect responders behavior that their models do not capture. Rather they argue that these other factors are likely to be of secondary importance and that inequality aversion is of primary importance. As such perhaps a more relevant question to ask is if either model is at least qualitatively consistent with the data; i.e., holding offer rates constant, are responders significantly more likely to accept offers under the positive consolation prize treatment and are they more likely to accept the larger the consolation prize?

Table 2 addresses these questions. All of the regressions are probits with the dependent variable taking on a value of 1 when the offer is accepted, and zero otherwise. They all employ a random effect error term, with subject as the random component. The first two probits serve as a baseline, showing the effects of the amount offered, and the square of the amount offered, on the likelihood of acceptance. Not surprisingly, in the first probit the OFFER coefficient is positive and significant at better than the 1% level, as higher offers increase the likelihood of acceptance. As the second probit indicates, higher offers are accepted with increasing probability as the square of the amount offered is also positive and significant at the 1% level. The next probit introduces a dummy variable, CDUM, that takes on a value of 1 in games with a positive consolation prize, and 0 otherwise. The coefficient value for CDUM is positive and significant at the 10% level indicating that there are marginally more acceptances, other things equal, in the positive consolation prize treatments. However, the introduction of separate dummies for each treatment shows that the probability of acceptance does *not* increase monotonically as the amount of the consolation prize increases, with the coefficient for the $12 treatment not significant at anything approaching conventional levels. The last two probits employ the dollar amount of the consolation prize as a continuous variable. The dollar amount offered

(CAMOUNT) by itself has a coefficient value that is essentially zero. CAMOUNT and CAMOUNTSQ (the dollar amount squared) are individually significant at very close to the 1% level, but opposite in sign with CAMOUNT positive and CAMOUNTSQ negative. This is totally at odds with the spirit of both models, and common sense. Perhaps the most charitable conclusion to reach on the basis of the results reported is that the existence of a positive consolation prize might increase the likelihood of acceptance, other things equal, but that the impact is not terribly robust as it has its smallest impact with the $12 treatment, precisely where its impact should be the largest. [6]

We can summarize the results of this section quite succinctly. The point predictions of both the BO and FS models are rejected as responders routinely reject offers that are less than half the social norm regardless of the fact that these rejections result in even greater income inequality relative to the non-responding player. The probits suggest that the existence of a positive consolation prize might have some marginal impact on the probability of acceptance, other things equal, but that this effect, if it is real, is not terribly robust as it fails to increase monotonically as the amount of the consolation prize increases.

Proposer Behavior:

Seventy-one percent of the 460 offers had equal amounts for Y and Z. Overall, the modal offer was at the equal distribution norm of ($5, $5, $5), with the median offer ($7, $4, $4). The amount that

[6] We tried other specifications as well. Dropping OFFERSQ and repeating the analysis the CDUM variable by itself is positive but small in value and not significant at anything approaching conventional levels (p > .40). Results for the other specifications remain essentially unchanged. We also tried specifications more consistent with the structure of the FS model. In some of these specifications we included amount offered to the other player along with amount offered to self. These regressions also provided no support for the positive consolation prize treatments materially increasing the probability of acceptance.

Xs kept for themselves decreased slightly over the course of the ten rounds under all treatments (see Figure 2 and Table 3).

[Insert Table 3 and Figure 2 here]

One point to note in Figure 2 and Table 3 is that Xs' keep is quite similar across treatments, beginning with the first round. Pairwise t-tests confirm the fact that there are no significant differences in the average keep value between any of the treatments in either the first or the last period of play ($p \geq .25$ in all cases, two-tailed tests). Most importantly, the average keep in the $12 consolation prize treatment is not noticeably different from the other treatments, so that few proposers thought to take advantage of the protection offered by any responder inequality aversion afforded by this large consolation prize. It winds up that proposers were correct on this point, but it does provide further, albeit indirect evidence, that the FS and BO models fail to organize the data very well.

Table 4 shows the rejection rates and expected values associated with different offers. Expected value drops rather sharply once offers are less than half the social norm (10, 2.50, 2.50), and when closest to the even distribution of (5, 5, 5). They are at a maximum between these values, with the modal distribution of (7, 4, 4) yielding nearly maximum expected value with relatively few rejections (11%).

[Insert Table 4]

Experiment 2:

The results of experiment 1 can be interpreted in one of two ways relative to the self-centered inequality aversion models of BO and FS: (1) The models are simply incorrect in how they capture behavior or (2) Responders frame their choices as one between themselves and X, largely ignoring Z in their deliberations. As such we decided to conduct an additional experiment to possibly tease apart these two alternatives.[7] To do this we introduced a negative consolation prize for Z in cases where Y rejected X's offer. Our thinking here was that the negative consolation prize might get responders to take fuller account of the impact of their decisions on Z. That is, its one thing to reject an offer that helps Z when exercising one's disdain for X's offer, its quite another to exercise it when it harms an innocent third party.[8] Of course, it is quite possible that most responders are willing to harm Z to get at X if X gets them angry enough. As such this experiment can only sort out between these two alternatives in cases where a negative consolation prize leads to fewer rejections, holding offers constant.

Procedures were the same as in experiment 1 with the following differences: All subjects received a starting capital balance of $15 in place of the $5 show up fee. In cases where Y turned down X's offer, Z *lost* $10. Further, the instructions were purposely couched in terms of Z *losing* $10 in cases where Y rejected X's offer (with losses subtracted from Z's starting capital balance). A total of 21 subjects arrived for this session.

Figure 3 reports the distribution of offers and rejections rates. The average rejection rate is down a bit relative to experiment 1, 14% versus average rejection rates of between 15% and 22% across treatments in experiment 1. Random effect probits testing for differences between experiment 1 and the negative consolation prize treatment are reported in Table 5. As expected, offers to Y (OFFER

[7] This experiment was suggested by Bob Slonim.

[8] In the choice under uncertainty literature there is a distinction between responses to gains versus losses (see, for example, Kahneman and Tversky, 1979). Although a different behavioral mechanism would be at work between this setting and the risky choice decisions, it is interesting to look for such effects in other contexts as well.

and OFFERSQ) play a critical role in responders' decisions to accept or reject offers. Further, contrary to our conjecture, the dummy variable distinguishing between experiment 1 and the present experiment, although positive, is not significant at conventional levels (p > .40). As such we conclude that changing the "consolation prize" so that it produces damage to an innocent third party, at least to the limited extent involved in this experiment, does not result in a dramatic increase in acceptance rates, other things equal.[9]

[Insert Figure 3 and Table 5 here]

Summary and Conclusions:

This paper provides a direct test of two recent models designed to explain breakdowns in the standard self-centered, utility maximizing model that underlies much of economic theory (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). Both models incorporate fairness considerations into some agents' utility functions in the form of concerns about the distribution of income between themselves and other players in the game. We test these models through a simple extension of the two-person ultimatum game, the game that motivates much of these modeling efforts. In our three person ultimatum game when responders reject an offer, a third player (the dummy) gets a positive consolation prize, which establishes income inequality between the responder and the "dummy" following rejections. This in turn should, according to the Bolton-Ockenfels model, eliminate *all* rejections, provided the offer is positive, and, according to the Fehr-Schmidt model, eliminate virtually all rejections with the $12

[9] These results do not appear to be terribly promising for the Charness and Rabin (2000) model. In that model, agents attempt to minimize the harm they do to the worst off agent in the reference group while balancing concerns for their own payoff, the total social surplus, and the intentions underlying other agents actions (reciprocity concerns). Although the Charness and Rabin model is consistent with our results if reciprocity concerns play an important enough role, if the do not (and there is little evidence to suggest that they matter to this extent), it can only be that

consolation prize. Although we did not expect the point predictions of the two models to be satisfied, we were unprepared for the limited (one might argue non-existent) drawing power of the underlying concept of inequality aversion when extended to this setting. Most devastating in this respect is the $12 consolation prize treatment which had no significant effect on rejection rates, even though these rejections establish strong income inequality between responders and the dummy. Further, the introduction of negative payoffs for the "dummy" does not induce responders to reject less often either.

What accounts for the near invisibility of the third player in terms of responders' reactions to X's offers? Although a complete answer to this question goes well beyond the scope of the present paper, there are a couple of obvious directions in which to look. One explanation, within the framework of the two models in question, is that there is some sort of framing effect, so that respondents do not account for the inequality resulting from their rejecting positive offers. As Fehr and Schmidt note "The determination of the relevant reference group and the relevant reference outcome for a given class of games is ultimately an empirical question." They go on to note, as we have assumed here, that the "natural reference group is simply the set of subjects playing against each other and that the reference point is equality." One way for both models to reconcile with the data reported here is to simply abandon this last assumption. However, this would be most unappealing, as it implies getting bogged down in ad hoc specifications of the relevant reference group for each setting, thereby robbing the models of two of their most attractive features, generality to a wide class of games and parsimony. Abandoning this assumption to explain our results is particularly unappealing since in many ways the three-person ultimatum game operationalized here is one of the simplest and most direct extensions of

---

people don't really care that much about third parties. And this is not really in their model.

these models to a new game that they were not originally designed to explain.

An obvious alternative explanation is that intentionality plays a much more important role than either model suggests. That is, income inequality resulting from an intentional action (the proposers low offer) is treated quite differently from unintentional income inequality (that between the responder and the dummy resulting from rejecting the offer). Both Bolton and Ockenfels and Fehr and Schmidt recognize that intentionality plays some role in the data their models are designed to explain (in particular, see the experiments reported in Blount, 1995; Charness, 1998; and Kagel, Kim and Moser, 1996). But they argue that these represent second order effects, not needed to explain the bulk of the data.[10] This, however, would not appear to be the case for our three-person game.

This concern for intentionality is captured in the reciprocity models offered in Rabin (1993) and Levine (1998). One outstanding criticism of these models is that they are not able to explain dictator game results very well, with many agents giving significant shares to total strangers who they can expect no reciprocal favors from. However, a paper by Oberholzer-Gee and Eichenberger (1998) suggests that these contributions may be largely artifactual, as giving in dictator games is sharply reduced if allocators are given an additional, but normally not preferred, choice.[11] This conclusion is reinforced by the results of the Guth and van Damme (1998) experiment, which essentially combined an ultimatum game with a dictator game, with the third agent, Z, playing the role of the passive agent in the dictator game. Two of the striking results from this experiment are the very small allocations to player Z, and the fact that not a single rejection could be attributed to the low share Z received.

---

[10] The small role attributed to intentionality in the Fehr-Schmidt model stands in marked contrast to Fehr's research on gift exchange in labor markets (see, for example, Fehr, Kirchsteiger, and Riedl (1993).

[11] One can also appeal to the common belief that "what goes around comes around" to explain giving in dictator games.

We would like to make one final point in closing. Although our experiment (and no doubt others) will reveal holes in the fairness models offered by Bolton and Ockenfels and Fehr and Schmidt, these shortcomings do not diminish their usefulness. Both have been able to systematically organize a large slice of existing experimental data (with minimal exclusion of embarrassing results) and suggest new ways of looking at the world. The fact that neither gets it quite right at first pass should be hardly surprising. The important thing is that they set the table for new ways of looking at the world and for conducting new and interesting experiments that will deepen our understanding of the important issue of how fairness considerations affect interactions between economic agents.

References

Blount, Sally. "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences." *Organizational Behavior and Human Decision Processes,* August 1995, *63*(2), pp. 131-144.

Bolton, Gary and Ockenfels, Axel. "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review,* March 2000, *90*, pp. 166-193.

Bolton, Gary and Ockenfels, Axel. "Strategy and Equity: An ERC analysis of the Güth-van Damme Game." *Journal of Mathematical Psychology,* June 1998, *42*(2), pp. 215-226.

Charness, Gary. "Attribution and Reciprocity in a Simulated Labor Market: An Experimental Investigation." Mimeographed, Universitat Pampea Fabra, 1998.

Charness, Gary and Rabin, Matthew, "Social Preferences: Some Simple Tests and a New Model," mimeographed, 2000.

Fehr, Ernst and Schmidt, Klaus. "A Theory of Fairness, Competition and Cooperation" *Quarterly Journal of Economics*, August 1999, *114* (3), pp. 769-816.

Fehr, E., Kirchsteiger, E., and Riedl, A., "Does Fairness Prevent Market Clearing? An Experimental Investigation." *The Quarterly Journal of Economics,* 1993, *108*, pp 437-459.

Forsythe, Robert; Horowitz, Joel; Savin, N. E. and Sefton, Martin. "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior*, 1994, *6*, pp. 347-369.

Güth, Werner and Van Damme, Eric. "Information, Strategic Behavior and Fairness in Ultimatum Bargaining: An Experimental Study." *Journal of Mathematical Psychology,* June 1998, *42*(2/3) pp. 227-247.

Kagel, John; Kim, Chung and Moser, Donald. "Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs. "*Games and Economic Behavior*, March 1996, *13*(1), pp. 100-110.

Kahneman, Daniel and Tversky, Amos. "Prospect Theory: An

Analysis of Decision Under Risk." *Econometrica,* March 1979, *47,* pp. 263-91.

Levine, David K. "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics,* July 1998, *1*(3), pp. 593-622.

Oberholzer-Gee, Felix and Eichenberger, Reiner. "Fairness! What Fairness? Focusing Effects in Dictator Game Experiments." Working paper, University of Zurich, 1998.

Rabin, Matthew. "Incorporation Fairness into Game Theory and Economics." *American Economic Review*, December 1993, *83*(5). pp. 1281-1302.

Roth, Alvin E.; Prasnikar, Vesna; Okuno-Fujiwara, Maeshiro and Zamir, Shmuel. A Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review*, December 1991, *81*(5), pp. 1068-1095.

Table 1: Rejection Rates by Consolation Prize Amount

Amount offered to responder	Rejection rates (# of rejections and # of offers made in parentheses)									
	Overall		C=$0		C=$1		C=$3		C=$12	
(0.10 , 0.74)	81%	(21/26)	83%	(10/12)	n/a	(0/0)	86%	(6/7)	71%	(5/7)
(0.75 , 1.24)	85%	(17/20)	88%	(7/8)	80%	(4/5)	n/a	(0/0)	86%	(6/7)
(1.25 , 1.74)	75%	(3/4)	n/a	(0/0)	50%	(1/2)	100%	(2/2)	n/a	(0/0)
(1.75 , 2.24)	76%	(16/21)	50%	(1/2)	80%	(4/5)	100%	(2/2)	75%	(9/12)
(2.25 , 2.74)	32%	(10/31)	50%	(1/2)	11%	(1/9)	40%	(8/20)	n/a	(0/0)
(2.75 , 3.24)	28%	(32/113)	47%	(7/15)	21%	(3/14)	23%	(7/31)	28%	(15/53)
(3.25 , 3.74)	22%	(15/68)	43%	(6/14)	33%	(1/3)	13%	(6/47)	50%	(2/4)
(3.75 , 4.24)	12%	(14/115)	23%	(6/26)	11%	(1/9)	5%	(1/22)	10%	(6/58)
(4.25 , 4.74)	6%	(2/35)	13%	(2/16)	0%	(0/8)	0%	(0/8)	0%	(0/3)
(4.75 , 5.24)	1%	(2/214)	1%	(1/93)	0%	(0/41)	3%	(1/40)	0%	(0/40)
(5. 25 +)	0%	(0/33)	0%	(0/12)	0%	(0/4)	0%	(0/1)	0%	(0/16)
Total	19%	(132/680)	21%	(41/200)	15%	(15/100)	18%	(33/180)	22%	(43/200)

Figure 1
Rejection rates by amount offered to responder

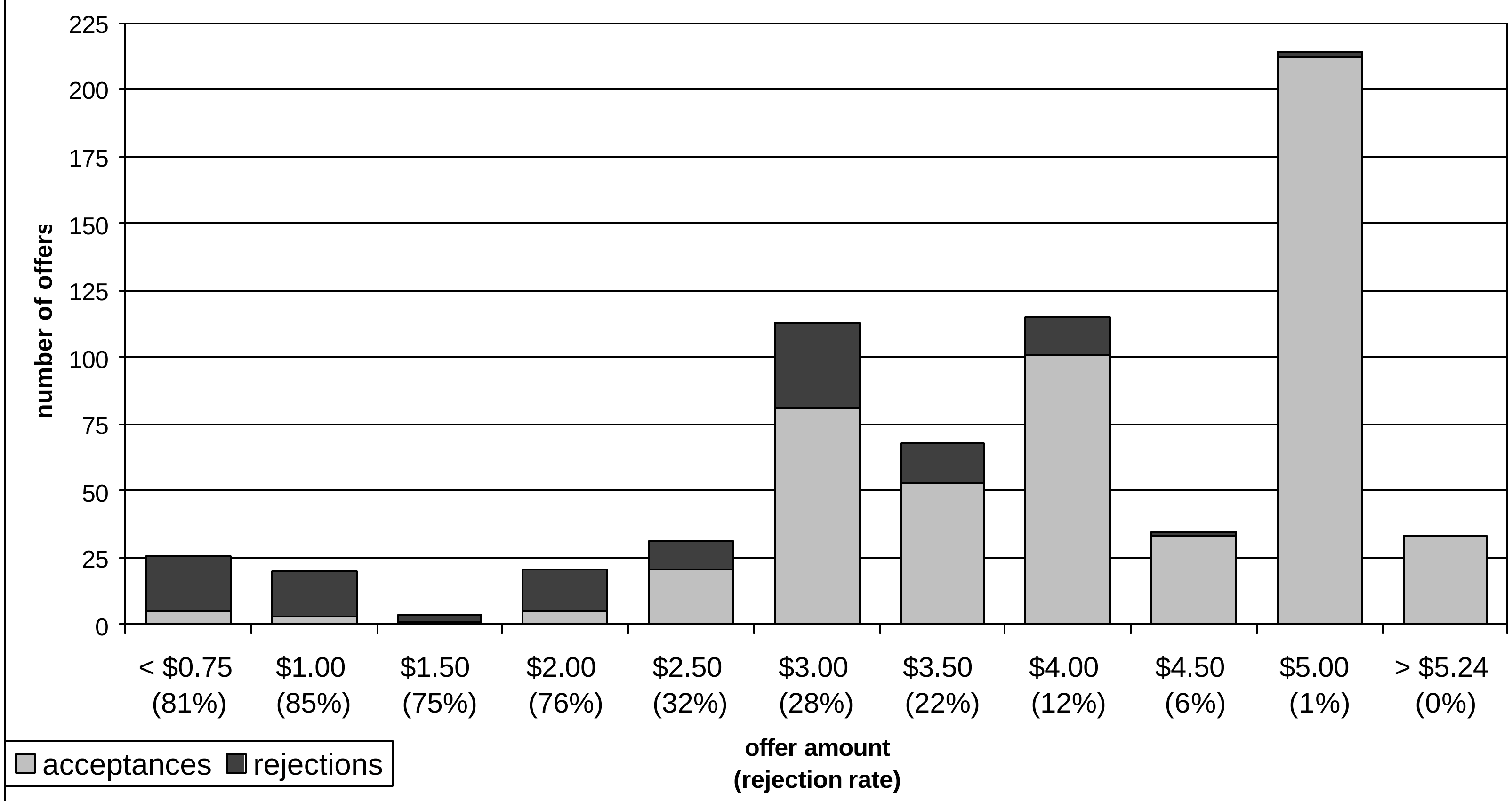

Table 2
Effects of Consolation Prize on Probability of Acceptance

	Log Likelihood
Accept = -2.75 + 1.21 OFFER (-0.348)** (0.12)**	-185.3
Accept = -1.99 + 0.52 OFFER + 0.12 OFFERSQ (0.44)* (0.30)+ (0.05)*	-182.3
Accept = -2.41 + 0.37 OFFER + 0.15 OFFERSQ + 0.71 CDUM (0.51)** (0.30) (0.06)** (0.40)+	-180.7
Accept = -2.36 + 0.32 OFFER + 0.16 OFFERSQ + 1.01 C1 + 1.07 C3 + 0.50 C12 (0.49)** (0.30) (0.06)** (0.56)+ (0.49)* (0.41)	-179.5
Accept = -1.92 + 0.52 OFFER + 0.12 OFFERSQ - 0.008 CAMOUNT (0.46)** (0.30)+ (0.05)* (0.028)	-182.3
Accept = -2.27 + 0.38 OFFER + 0.15 OFFERSQ + 0.46 CAMOUNT - 0.04 CAMOUNTSQ (0.47)** (0.29) (0.05)** (0.19)* (0.01)*	-180.1

**Significantly different from 0 at better than the 1% level.
* Significantly different from 0 at better than the 5% level.
+ Significantly different from 0 at better than the 10% level.

OFFER = Amount offered
OFFERSQ = OFFER squared
CDUM = 1 if positive consolation prize, = 0 otherwise
C1 = 1 if $1.00 consolation prize, = 0 otherwise
C3 = 1 if $3.00 consolation prize, = 0 otherwise
C12 = 1 if $12.00 consolation prize, = 0 otherwise
CAMOUNT = amount of consolation prize (as a continuous variable)
CAMOUNTSQ = CAMOUNT squared

78 subjects, 680 observations

Table 3
Keep for Proposer - Mean amount by round

round	c=0	c=1	c=3	c=12	average
1	7.06	7.10	8.20	8.18	7.62
2	7.07	7.15	8.20	8.01	7.59
3	6.45	6.80	7.00	7.70	6.99
4	6.40	6.63	7.82	7.35	7.03
5	6.60	6.63	7.61	8.50	7.33
6	6.60	6.53	7.91	7.01	6.99
7	6.55	7.05	7.93	7.20	7.16
8	6.88	7.75	7.93	6.95	7.36
9	6.45	7.65	8.10	7.35	7.37
10	6.75	6.78	7.79	7.01	7.06

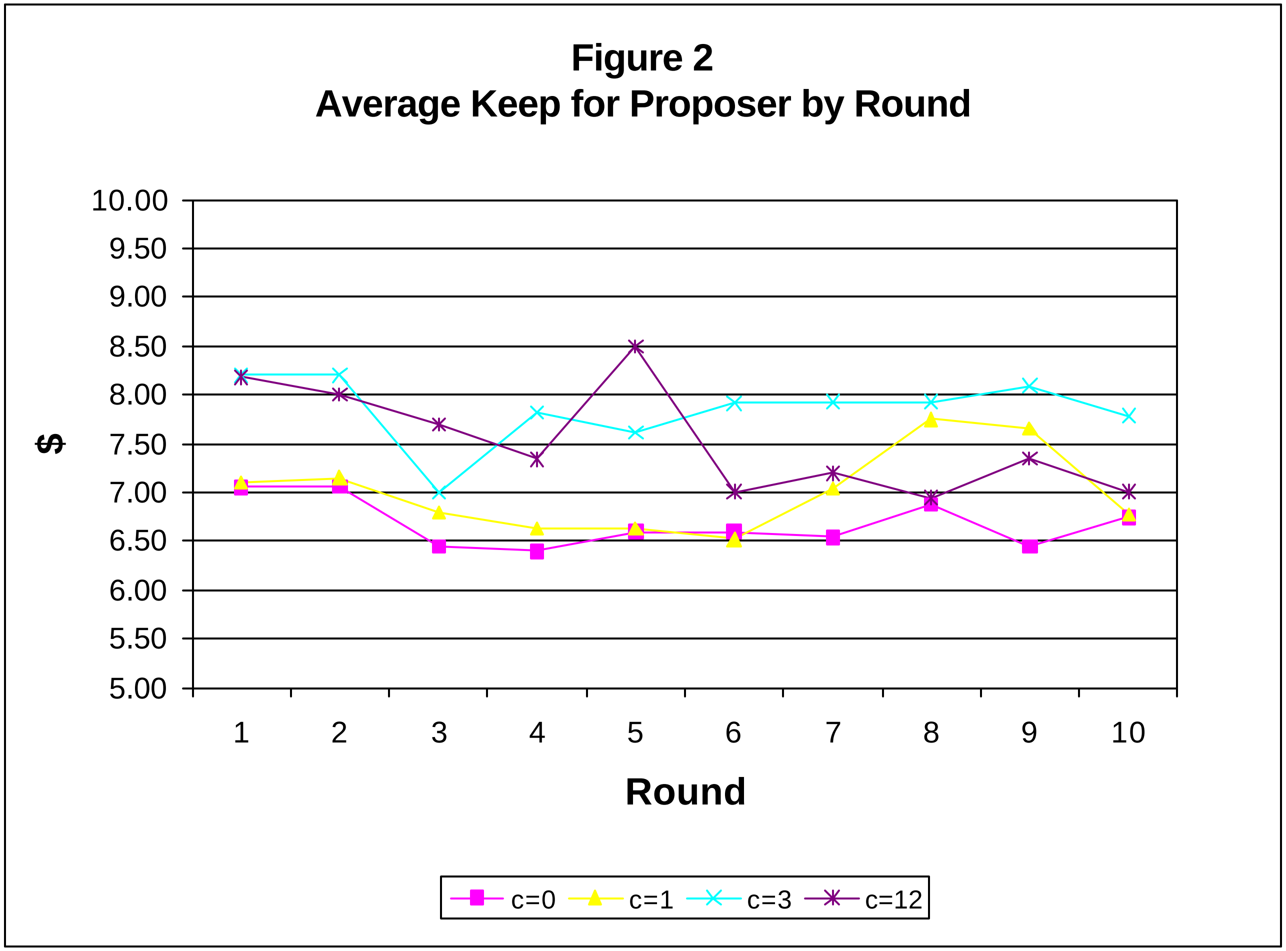
Figure 2
Average Keep for Proposer by Round
$
10.00
9.50
9.00
8.50
8.00
7.50
7.00
6.50
6.00
5.50
5.00
1
2
3
4
5
6
7
8
9
10
Round
c=0
c=1
c=3
c=12

Table 4:
Proposer's Expected Value Calculations

Offer	Rejection rate	Expected Value to Proposer
$13.00, $1.00, $1.00	85%	$1.95
$12.00, $1.50, $1.50	75%	$2.75
$11.00, $2.00, $2.00	67%	$3.67
$10.00, $2.50, $2.50	36%	$6.44
$9.00, $3.00, $3.00	30%	$6.26
$8.00, $3.50, $3.50	22%	$6.24
$7.00, $4.00, $4.00	11%	$6.24
$6.00, $4.50, $4.50	6%	$5.67
$5.00, $5.00, $5.00	1%	$4.96

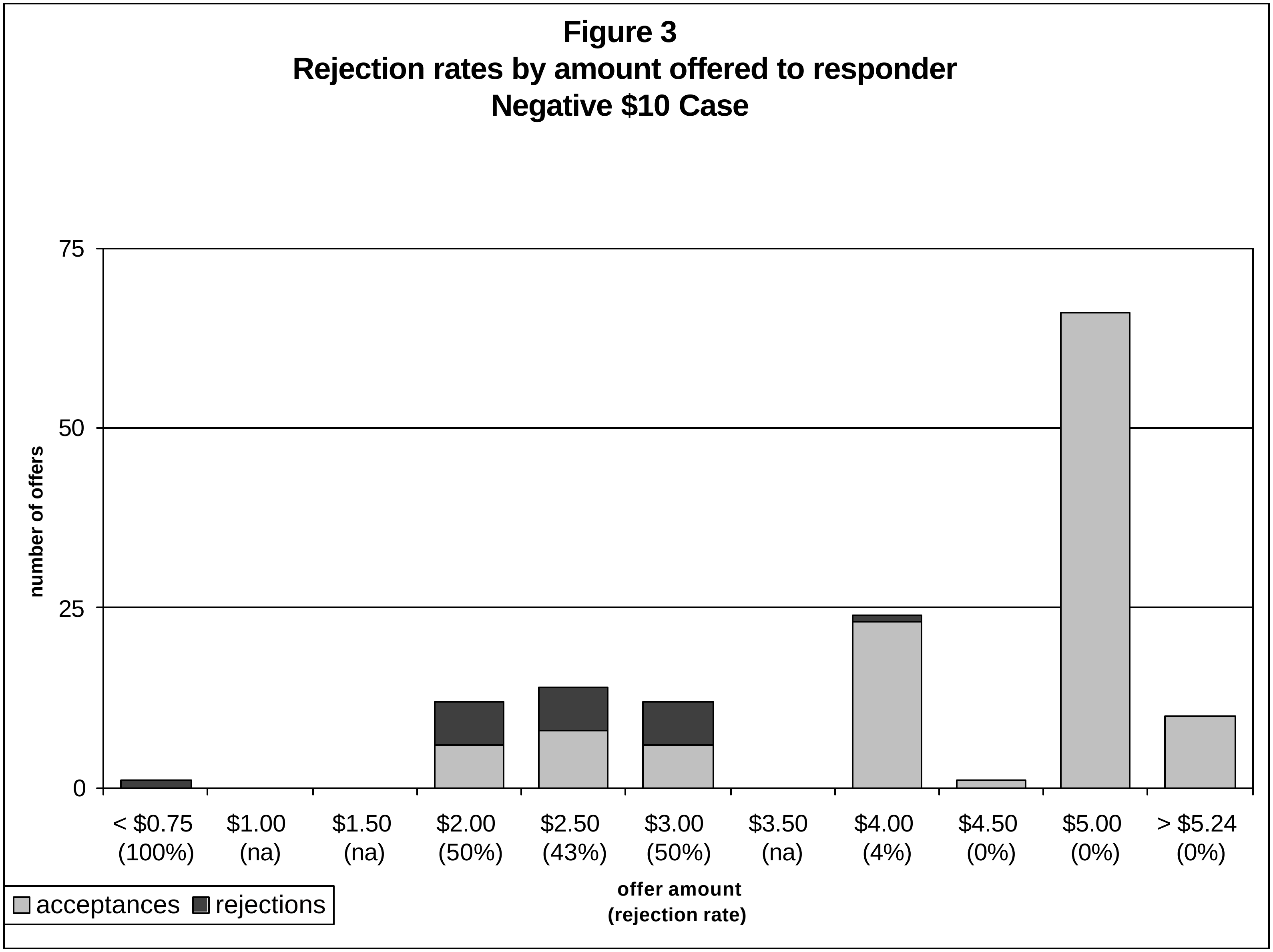
Figure 3
Rejection rates by amount offered to responder
Negative $10 Case
number of offers
75
50
25
0
< $0.75 (100%)
$1.00 (na)
$1.50 (na)
$2.00 (50%)
$2.50 (43%)
$3.00 (50%)
$3.50 (na)
$4.00 (4%)
$4.50 (0%)
$5.00 (0%)
> $5.24 (0%)
offer amount
(rejection rate)
acceptances
rejections

Table 5
Random Effects Probit for Negative Consolation Prize

	Log Likelihood
Accept = -1.95 + 0.43 OFFER + 0.14 OFFERSQ + 0.40 DNEG10 (0.43)** (0.29) (0.05)** (0.47)	-209.8

**Significantly different from 0 at better than the 1% level.

OFFER = Amount offered
OFFERSQ = OFFER squared
DNEG10 = 1 if consolation prize is -$10, = 0 otherwise

92 subjects, 820 observations