HYPOTHETICAL AND REAL INCENTIVES IN THE ULTIMATUM GAME AND ANDREONI'S PUBLIC GOODS GAME:

AN EXPERIMENTAL STUDY

Mark T. Gillis
West Virginia University

and

Paul L. Hettler, Ph.D.
California University of Pennsylvania

INTRODUCTION

> The reliance on hypothetical choices raises obvious questions regarding the validity of the method and the generalizability of the results. By default, the method of hypothetical choices emerges as the simplest procedure by which a large number of theoretical questions can be investigated. The use of the method relies on the assumption that people often know how they would behave in actual situations of choice, and on the further assumption that the subjects have no special reason to disguise their true preferences.
>
> —Kahneman and Tversky [1979, 265]

Based on rational self-interest, many economists believe that only real financial incentives will elicit true economic behavior from participants in economic experiments. Thus, the use of hypothetical incentives is generally restricted to classroom examples (and even in this case, extra credit points or some other academic reward may be used). The idea that people will show their real preferences in response to hypothetical incentives is believed by many to be unrealistic. It is widely accepted that while hypothetical examples can help in explaining economic situations and gauging how people might react to a situation, they will not provide an exact replication of human economic behavior.

However, there is little actual experimental evidence to suggest that hypothetical incentives do not give similar results compared to financial incentives in at least some economic experiments. Some experiments have been conducted using hypothetical incentives, relying on the assumption that people have no reason not to act in an economically rational manner regardless of the type of incentive. For many experimental

Paul L. Hettler: Department of Business and Economics, California University of Pennsylvania, Box 74, 250 University Ave, California, Pennsylvania 15419. E-mail: hettler@cup.edu.

situations there is little reason to believe that people will not act in such a manner, even if incentives are hypothetical. The use of financial incentives is one of the few features that distinguishes economics experiments from those in other disciplines. As Camerer and Hogarth [1999, 7] point out:

> Economists presume that experimental subjects do not work for free and work harder, more persistently, and more effectively, if they earn more money for better performance. Psychologists believe that intrinsic motivation is usually high enough to produce steady effort even in the absence of financial rewards; and while more money might induce more effort, the effort does not always improve performance, especially if good performance requires subjects to induce spontaneously a principle of rational choice or judgment.

Given the high costs that are associated with conducting economic experiments, finding more cost effective methods has become a very important issue in experimental economics. Experimenters have adopted several methods of reducing the financial cost of experimentation ranging from using hypothetical incentives to using random-pay methods to conducting experiments in foreign countries. Understanding the impact of hypothetical incentives on experimental outcomes is crucial. If experimental players behave in the same manner with hypothetical incentives as they do with real incentives then not only can enormous amounts of money be saved, but a wider variety of experiments can be conducted.

This paper examines how subjects in the Ultimatum Game (UG) and Andreoni's [1993] Public Goods Game (APGG) behave in response to hypothetical and real incentives. The results from these two experiments are used to determine if economic games played with hypothetical incentives give accurate results. Several explanations exist as to why results may or may not be the same in different games; these are discussed in the conclusion.

The paper also uses the experimental results from these two games to further investigate how players make decisions in economic experiments. Economic theory usually points to self-interest as the main motivation in economic game play. However, in practice results do not always agree with this, thus we also look at how other aspects of behavior, such as fairness, complexity, competitiveness, and sequentiality and order, can affect game play. The significance of each of these is also important in determining whether hypothetical and real incentives can give similar results. The remainder of the paper is organized as follows: the next section reviews the literature and summarizes past experimental results from UG and APGG; the one following that presents the experimental design; then there is a section that describes the results, and finally a concluding section which suggests future research.

LITERATURE REVIEW

In order to increase the size and scope of their experiments, experimenters have been using a wide range of methods to lower the overall costs of experimentation.

These methods have allowed economists not only to test a wider range of economic theories at a lower cost, but also to conduct experiments that may not be feasible under other conditions.

Bolle [1990] suggests a method of reducing experimentation costs in which experimenters simply pay a random player, or random set of players, from the game. In this situation, he suggests that at the end of the experiment, a subset of players, chosen at random, be paid.[1] Thus, experimenters could both lower the overall costs of experiments and significantly raise the stakes in experiments without significantly raising costs.[2] Bolle provides theoretical and experimental evidence from UG that such a method works. Sefton [1992] provides evidence from the Dictator Game[3] that suggests that Bolle's approach may not always work, finding that subjects are more generous when paid for randomly selected rounds. Sefton suggests that the variation in these results may be caused by the type of game or by the size of the stakes in the game.

Other experimenters have conducted experiments in less developed countries in order to gain experimental evidence about "high-stakes" games at relatively low cost. Cameron [1999] conducted a UG experiment in Indonesia that allowed for stakes to be raised to three times the average monthly expenditures of the participants. Slonim and Roth [1998] conducted a similar experiment in the Slovak Republic, which allowed the stakes to be raised to the equivalent of 62.5 hours of average wage. On the surface, the results from this approach are very effective for lowering costs and expanding the range of economic experiments. However, comparing results from such studies implies that ethnic and cultural backgrounds do not affect bargaining/economic behavior. Roth, et.al., [1991] investigate this very fact as they compare the results of UG and a market auction game across four different countries. They find that there are some important differences in market behavior that may be attributed to cultural differences. Thus, relying on high stakes games from other countries alone provides insufficient, or at least incomplete, results.

Blumenschein, et. al. [1997] investigate how hypothetical and real payments differ in Vickrey auctions.[4] They find that bids in Vickrey auctions involving hypothetical payments are significantly higher than bids in such auctions involving real payments. Blumenschein, et. al. conclude that there is a bias in any study that involves hypothetical payments. However, they do not present reasons for why this bias exists, nor do they give justifying explanations as to why it should hold for hypothetical incentives in all economic experiments. While the results found by Blumenschein, et. al. do suggest that hypothetical incentives may not work in economic experiments, they only consider one type of experiment.

Kahneman and Tversky [1979] rely solely on hypothetical incentives in their study of risk aversion. They suggest that their study depends on the assumptions that people know how they will act in actual situations and that they have no reason to hide their true preferences. They further argue that since any economic study has some drawbacks, relying on hypothetical choices is not any more detrimental. The study by Kahneman and Tversky has since been challenged by Holt and Laury [2002], who find that subjects become increasingly risk averse as payoffs increase with real incentives, whereas increasing the scale of hypothetical incentives has no effect on behavior. Harrison, et. al. [2005] and Holt and Laury [2005] confirm this result even after correcting for possible order effects.

Thaler [1986] addresses the issue of whether small incentives, relative to larger ones, are efficient in economic games. Thaler argues that there is not enough evidence to show that raising stakes is more likely to elicit economic behavior. This further suggests that if higher stakes do not guarantee better game play, then real incentives may not be needed at all.

Forsythe, et. al. [1994] take a different approach to economic theory in using hypothetical incentives in economic experiments. Although the monetary significance of hypothetical incentives is in itself very important, Forsythe, et. al. present the idea that hypothetical incentives can be used to further test what players base their game play on. They use hypothetical payments to both test whether players behave similarly in Ultimatum and Dictator games with hypothetical and real payoffs and to determine what role fairness plays in bargaining games. They determine that fairness is not as significant a factor as many have previously thought.

Although there have been several attempts to determine ways of lowering costs in economics experiments and to investigate how hypothetical incentives change a player's game play, no solid conclusions have been drawn and even fewer reasons have been presented which explain observed outcomes. In this paper, we further investigate the relationship between hypothetical incentives and behavior in economics experiments and present possible explanations for differences or anomalies in the results. The varying results found tend to support the aggregate inconclusiveness of previous studies about hypothetical incentives.

Andreoni's Public Goods Game

APGG was first used by Andreoni [1993] to test the public goods crowding out hypothesis, that is, involuntary contributions (taxes) will crowd out voluntary contributions to public goods. Andreoni uses a payoff matrix[5] that involves "token" contributions to a group investment from three separate players. These contributions determine a player's individual payoff. The three players in Andreoni's Game are assigned to a random group, in which they do not know the identities or individual contributions of the other players in their group. The matrix reflects a public good in that as one player's contribution rises, the payoff for the other players rises. In Andreoni's Game, the Nash equilibrium is for each subject to contribute three tokens and the Pareto efficient outcome is for each subject to contribute six tokens.

APGG, in its original play, involves two stages. The first is based on the original payoff matrix, where three players can contribute from 0 to 7 tokens. The second stage involves a mandatory contribution of at least 2 tokens, and then allows players to contribute anywhere from 0 to 5 additional tokens. The payoff matrix for this part is the same as the original payoff matrix, but with the first two columns and four rows removed. This mandatory contribution plays the role of a tax towards the public good. In his original study, Andreoni finds that the average total contribution under the taxation situation is significantly larger than the average contribution without the tax. Furthermore, Andreoni finds crowding out that is, although still incomplete, substantially larger than the amounts found in econometric studies.

For this paper, only the first stage of Andreoni's experiment will be compared with hypothetical and real incentives. Using real incentives, Andreoni reported an average contribution of 2.78 tokens for the first stage.

Ultimatum Game

The UG is a simple two-stage bargaining game involving a proposer and a responder. In the game, the proposer makes a decision to split some pot into an amount he wants to keep for himself and an amount he offers to the responder. The responder is then given the choice of accepting the offer or rejecting the offer, in which case neither player receives anything. Economic theory suggests that the proposer should always offer the minimum amount to the responder, since both players would be better off with anything than with nothing. However, historical evidence differs significantly as splits typically follow a 60-40 average, and responders rarely are willing to accept any offer [Slonim and Roth, 1998; Camerer and Thaler, 1995; Roth, et. al. 1991].

Experimenters generally believe that the decision made by the proposer is based on some combination of three factors: self-interest, fairness, and fear of rejection. Self-interest refers to a players concern only for maximizing his own payoff. A player concerned only with self-interest would choose to offer the Nash equilibrium of 0 to the second player.[6] Fairness refers to a player wanting to make a fair offer to the responder. The ideal offer in this case is believed to be an equal 50-50 split between the two players. This means that if fairness is a factor in decision-making in the UG, then it has a negative effect on the Nash equilibrium point.[7] The final possible factor is the fear of rejection of an offer. This would also have a negative effect on the Nash equilibrium and, on its own, encourages a player to offer as much as possible to the responder.[8]

Given the amount of evidence showing that players do not rationally play the game based on self-interest alone, it is generally assumed that a proposer in the UG makes his decision based on a combination of self-interest and fairness, or self-interest and fear of rejection.[9]

EXPERIMENTAL DESIGN

Four sessions were conducted as part of this experiment. All participants were students enrolled in principles of economics courses at Duquesne University or California University of Pennsylvania. In exchange for their participation, students received some pre-determined amount of extra-credit in their course. In each session participants played both the UG and APGG. Table 1 summarizes the order of play and incentive type for each session. All participants entered the sessions unaware that they may be paid for their participation. For this reason, the games with real incentives were always played second.[10]

At the beginning of each experiment, subjects were told, either to "do their best" or that they would be paid, in cash, based on their participation at the end of the game. The UG was played with a $10 pot. Subjects were randomly assigned the role of proposer or responder as they entered the room. Proposers were asked to make any

offer from $0 to $10, down to the nearest nickel, to a randomly selected responder. Decision forms were collected by the experimenters and randomly assigned to responders. Responders were then asked to decide whether to accept the offer, with both players receiving their share of the split, or to reject the offer in which case both the proposer and the responder would receive nothing. After responders made their decisions, the forms were re-collected and re-distributed to the proposers. All participants played 10 rounds (although they were unaware of how many rounds would be played). For the real incentive session, subjects were paid in cash for one randomly selected round at the end of the session.[11]

TABLE 1
Games Played and Incentive Types, by Session

Session	First Game/Incentive Type	Second Game/Incentive Type
1	APG Game Hypothetical	Ultimatum Game Real
2	Ultimatum Game Hypothetical	APG Game Real
3	APG Game Hypothetical	Ultimatum Game Hypothetical
4	Ultimatum Game Hypothetical	APG Game Hypothetical

Subjects in APGG were asked to choose from 0 to 7 tokens to contribute to the group investment. After players denoted how many tokens they wished to contribute, decision forms were collected by the experimenters and randomly assigned to a group. Experimenters noted what group member each player was in and how many tokens the other group members contributed and returned the forms to the participants. Subjects then calculated their payoffs according to the payoff matrix and began making their decisions for the next round. For APGG subjects played a randomly selected number of rounds, between 6 and 10. For the real incentive session, subjects were paid in cash the sum of their earnings from all rounds of play (in cents).[12]

DATA ANALYSIS

Ultimatum Game

Table 2 summarizes the results from the UG, by session. Data from Slonim and Roth's [1998] low stakes session is included as an additional real incentive comparison. The upper half of the table presents first round results, while the lower half summarizes all rounds.[13] Session 3 stands out as having higher offers; Session 4 has a much higher first round acceptance rate. Comparing the first round results to all rounds, there is a clear difference between the real and hypothetical incentive sessions. While the average amount offered remains relatively stable in the real incentive session, the average amount offered falls significantly in each of the hypothetical sessions (an average of 7 percent). In all but Session 4, the mean acceptance rate rises after the first round; however it generally rises less in the hypothetical sessions.

Figure 1 and Figure 2 present histograms of the offers made in the first round and all rounds, respectively. Session 3 again seems to stand out, especially in the first round. When examining all rounds, the distribution of offers in the real incentive session (Session 1) appears to differ from the other sessions.[14]

TABLE 2
Ultimatum Game Results Summary
Ultimatum Game Results (first round)

| Session | Incentive Type | Observations | Amount Offered | | | Mean Acceptance |
			Mean	Median	Std. Dev.	
1	R	12	3.979	4.375	1.325	0.250
2	H	12	3.358	4.00	2.033	0.250
3	H	8	4.625	5.00	0.744	0.125
4	H	10	3.300	3.75	2.091	0.700
All hypothetical	H	30	3.677	4.00	1.836	0.367
Slonim and Roth (1998) 'low stakes' session:						
Practice Round	H	24	4.988	5.00	0.782	0.958
First Round	R	24	4.538	4.50	0.487	0.875

Ultimatum Game Results (all rounds)

| Session | Incentive Type | Observations | Amount Offered | | | Mean Acceptance |
			Mean	Median	Std. Dev.	
1	R	120	3.998	4.00	1.153	0.783
2	H	120	2.961	3.00	1.515	0.583
3	H	80	4.091	4.00	0.996	0.500
4	H	100	3.053	3.375	1.597	0.690
All hypothetical	H	300	3.413	3.75	1.440	0.550
Slonim and Roth (1998) 'low stakes' session:						
	R	240	4.450	4.70	0.693	0.825

FIGURE 1
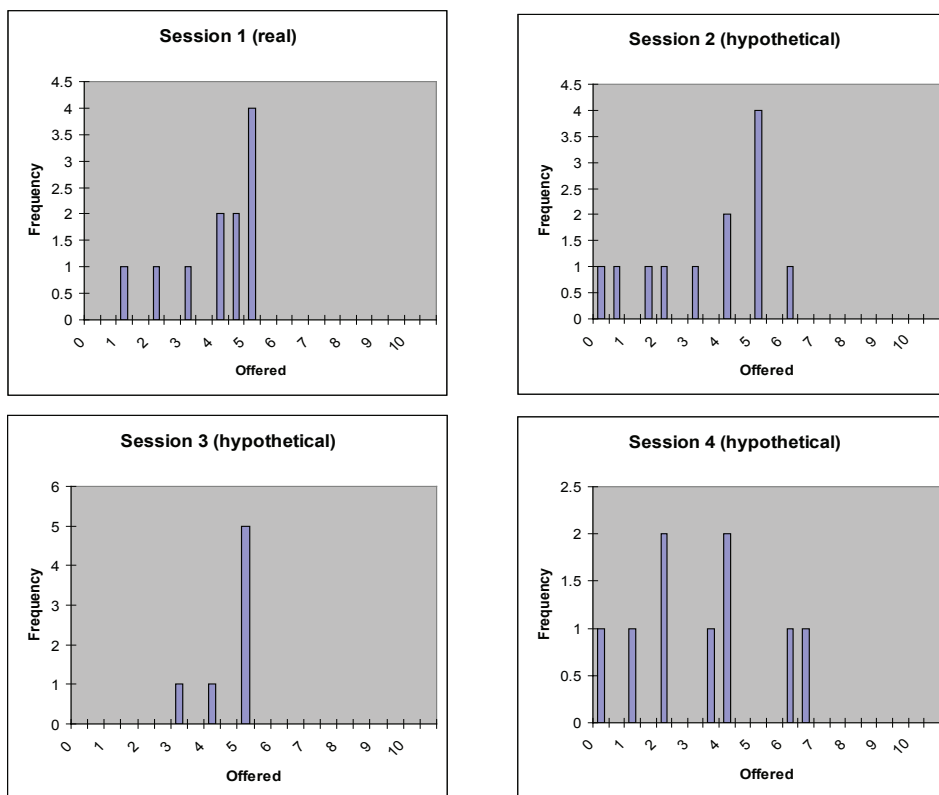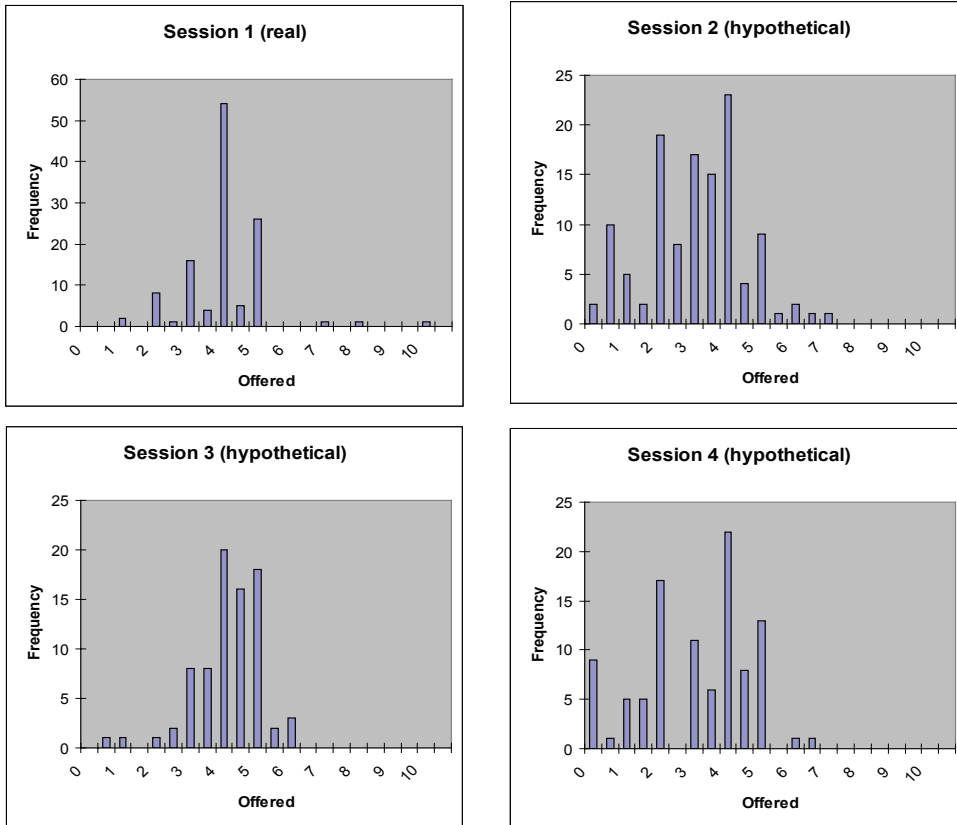### Histograms of First Round Offers in the Ultimatum Game

FIGURE 2
Histograms of Offers in the Ultimatum Game (all rounds)

The results of *t*-tests for the equality of means are presented in Table 3 for the first round and for all rounds. The shaded cells indicate tests of the real versus hypothetical incentive hypothesis. The *t*-tests indicate that there is no statistical difference between the first round offers in our real incentive session and any of our hypothetical incentive sessions. However, the average offer in all rounds of Sessions 2 and 4 are statistically significantly different from the average real incentive session offer at the 1 percent level. Players in both of those sessions offered less than players in the real incentive session (they were playing closer to the Nash equilibrium).

By aggregating the data across all rounds, these *t*-tests assume that each decision in each round is an independent observation; this is clearly not the case as participants may be choosing strategies in which choices in later rounds are predicated on behavior observed in earlier rounds. In the extreme this might imply that each session is an independent observation; however, since our variable of interest does not vary within a session such a restriction is not practical. Here we apply the individual-specific random effects model to perform our hypothesis tests. Table 4 presents these regression results. The same differences noted with the *t*-tests above can be seen here as well. The

real incentive variable has a statistically significant positive impact on amount offered when examining data from all rounds, but is insignificant in the first round data.

TABLE 3A
Ultimatum Game: t-tests for Equality of Means, by Session (Offered)

Session		1 Real	2 Hypothetical	3 Hypothetical	4 Hypothetical
	Mean	3.979 (1st)	3.358 (1st)	4.625 (1st)	3.300 (1st)
		3.998 (all)	2.961 (all)	4.091 (all)	3.053 (all)
1	3.979 (1st)		-0.621	-0.646	0.679
Real	3.998 (all)		(-0.882)	(-1.391)	(0.895)
2	3.358 (1st)	-1.037		-1.267	0.058
Hypothetical	2.961 (all)	(-5.969)***		(-1.970)*	(0.066)
3	4.625 (1st)	0.092	-1.130		1.325
Hypothetical	4.091 (all)	(0.602)	(-6.364)***		(1.877)*
4	3.300 (1st)	-0.945	-0.092	1.037	
Hypothetical	3.053 (all)	(-4.941)***	(-0.436)	(5.329)***	

Numbers in italics refer to first round data. First round t-tests appear above the diagonal. Cells in the body of the table contain the difference between the means and related t-statistics (in parentheses). * = 10% significance level, ** = 5% significance level, *** = 1% significance level. Shaded cells present tests of the hypothetical pay hypothesis. All t-tests are based on unequal variances.

TABLE 3B
Ultimatum Game: t-tests for Equality of Means, by Incentive (Offered)

Session		Session 1 Real	
	Mean	3.998 (all)	3.979 (1st)
2, 3,	3.677 (1st)		-0.303
and 4			(-0.595)
Hypothetical	3.293 (all)	-0.706	
		(-5.174)***	

Numbers in italics refer to first round data. First round t-tests appear above the diagonal. Cells in the body of the table contain the difference between the means and related t-statistics (in parentheses). * = 10% significance level, ** = 5% significance level, *** = 1% significance level. All t-tests are based on unequal variances.

TABLE 4
Ultimatum Game: Random Effects Regression,
Dependent Variable Amount Offered

	All Rounds	First Round Only
Incentive (=1 if real)	0.7055	0.3025
	(1.97)**	(0.6)
Round	-0.0126	—
	(-0.72)	
Constant	3.362	3.677
	-15.73***	-10.89***
R-squared	0.05	0.01
Observations	420	42

Cells in the body of the table present the coefficient estimate and z-statistic (in parentheses). * = 10% significance level, ** = 5% significance level, *** = 1% significance level

Finally, random effects probit regressions are used to compare whether or not hypothetical versus real incentives affect acceptance rates in the UG.[15] The results in Table 5 indicate that while amount offered has a statistically significant impact on acceptance, the coefficient on the incentive variable is insignificant.

TABLE 5
Ultimatum Game: Random Effects Probit Model,
Dependent Variable Accepted (=1 if Proposer's offer is accepted)

	All Rounds	First Round Only
Incentive (=1 if real)	0.2072	-0.8593
	(0.71)	(-1.02)
Amount Offered	0.8364	1.043
	(7.59)***	(2.42)**
Round	0.003	--
	(0.11)	
Constant	-2.405	-2.279
	(-5.95)***	(-2.08)**
Log Likelihood	-193.2	-8.834
Observations	420	42

Cells in the body of the table present the coefficient estimate and z-statistic (in parentheses). * = 10% significance level, ** = 5% significance level, *** = 1% significance level.

Generally speaking, the results from the Ultimatum game are inconclusive. Sessions 2 and 4 had similar mean offers, but very different acceptance rates. Sessions 1 (real) and 3 also had similar means (but different from the other two sessions), but the acceptance rate in Session 1 was higher. While generally results from all rounds indicate a difference between the results of the real and hypothetical sessions, results from the first round of play do not indicate such a difference.

Andreoni's Public Goods Game

Table 6 presents summary statistics for the APGG by session. Data from Andreoni [1993] is provided as an additional real incentive comparison. The upper half of the table presents first round results, while the lower half summarizes all rounds. Session 3 again stands out; the average contribution and payoff were higher in the first and all rounds. Beyond this, there appears to be no differences between sessions.

Figures 3 and 4 present histograms of the number of tokens contributed in each session for APGG for the first and all rounds, respectively. Overall, these distributions appear very similar.[16] Comparing the first round results to all rounds, it appears that the amount contributed decreases more in the hypothetical sessions than in the real incentive session.

The results of *t*-tests for the equality of means for the number of tokens contributed in APGG are presented in Table 7 for the first round and for all rounds. The shaded cells indicate tests of the real versus hypothetical incentive hypothesis. The *t*-tests indicate that there is no statistical difference between the offers in our real incentive session and those seen in Sessions 1 or 4 (first round and all rounds); however, we find a significant difference between the real session and Session 3. Players in that session offered slightly more than the Nash equilibrium 3 tokens, while players in the other sessions offered slightly less.

TABLE 6
Andreoni's Public Goods Game Results Summary

Andreoni's Public Goods Game Results (first round)

Session	Incentive Type	Observations	Amount Contributed Mean	Median	Std. Dev.	Mean Payoff
1	H	23	2.913	3.00	1.703	57.783
2	R	23	2.695	2.00	1.550	55.304
3	H	16	3.563	3.00	1.672	76.500
4	H	18	2.889	3.00	1.231	58.444
All hypothetical	H	57	3.088	3.00	1.562	63.246
Andreoni (1993)						
First Round (collected 1989)	R	18	2.83	3.00	1.150	62.333
First Round (collected 1990)	R	18	2.94	3.00	1.552	60.778

Andreoni's Public Goods Game Results (all rounds)

Session	Incentive Type	Observations	Amount Contributed Mean	Median	Std. Dev.	Mean Payoff
1	H	184	2.745	2.50	2.063	55.832
2	R	138	2.645	3.00	1.894	53.283
3	H	160	3.206	3.00	2.089	67.744
4	H	126	2.730	3.00	1.536	54.595
All hypothetical	H	470	2.898	3.00	1.954	59.325
Andreoni (1993)	R	720	2.788	3.00	1.561	57.829

FIGURE 3
Histograms of Amount Contributed in Andreoni's Public Goods Game (first round)
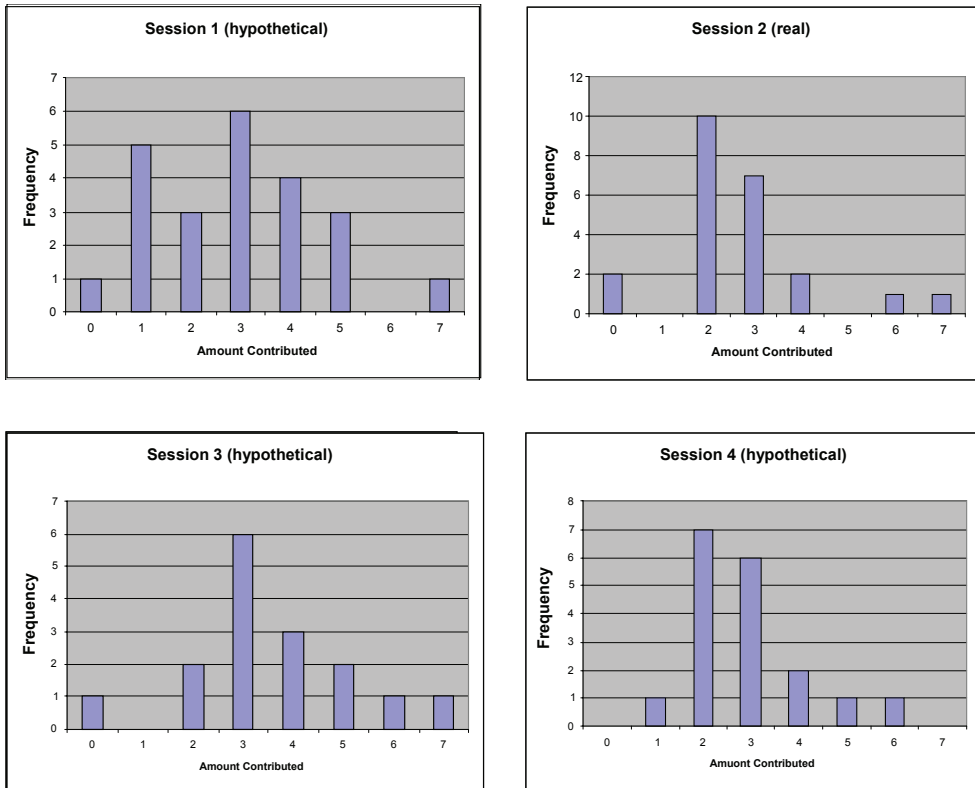
FIGURE 4

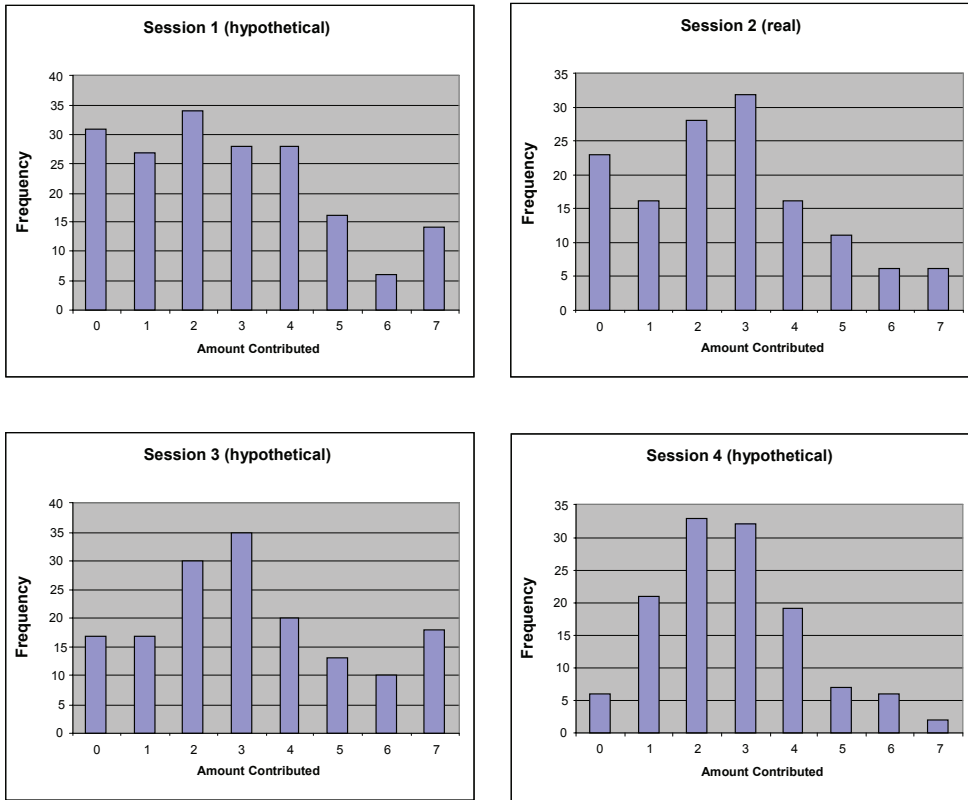**Histograms of Amount Contributed in Andreoni's Public Goods Game (all rounds)**

TABLE 7A

Andreoni's Public Goods Game: t-tests for Equality of Means,
by Session (Amount Contributed)

Session		1 Hypothetical	2 Real	3 Hypothetical	4 Hypothetical
	Mean	*2.913 (1st)*	*2.696 (1st)*	*3.562 (1st)*	*2.889 (1st)*
		2.744 (all)	2.645 (all)	3.201 (all)	2.730 (all)
1	*2.913 (1st)*		*-0.217*	*-0.649*	*0.241*
Hypothetical	2.744 (all)		*(-0.453)*	*(-1.184)*	*(0.053)*
2	*2.696 (1st)*	-0.099		*-1.267*	*-0.193*
Real	2.645 (all)	(-0.450)		*(-1.970)*	*(-0.445)*
3	*3.562 (1st)*	-0.462	-0.561		*0.674*
Hypothetical	3.201 (all)	(-2.056)**	(-2.432)**		*(1.324)*
4	*2.889 (1st)*	0.014	-0.085	0.476	
Hypothetical	2.730 (all)	(0.070)	(-0.403)	(2.220)**	

Numbers in italics refer to first round data. First round t-tests appear above the diagonal. Cells in the body of the table contain the difference between the means and related t-statistics (in parentheses). * = 10% significance level, ** = 5% significance level, *** = 1% significance level. Shaded cells present tests of the hypothetical pay hypothesis. All t-tests are based on unequal variances.

TABLE 7B
Andreoni's Public Goods Game: t-tests for Equality of Means,
by Incentive (Amount Contributed)

Session		Session 2 Real	
	Mean	2.645 (all)	*2.695 (1st)*
1, 3, and 4	*3.088 (1st)*		*0.392*
			(1.022)
Hypothetical	2.899 (all)	0.253	
		(1.369)	

Numbers in italics refer to first round data. First round t-tests appear above the diagonal. Cells in the body of the table contain the difference between the means and related t-statistics (in parentheses). * = 10% significance level, ** = 5% significance level, *** = 1% significance level. All t-tests are based on unequal variances.

TABLE 8
Andreoni's Public Goods Game: Random Effects Regression,
Dependent Variable Amount Contributed

	All Rounds
Incentive (=1 if real)	-0.3143
	(-1.03)
Round	-0.0734
	(-2.57)***
Constant	3.216
	(15.580)***
R-squared	0.01
Observations	608

Cells in the body of the table present the coefficient estimate and z-statistic (in parentheses). * = 10% significance level, ** = 5% significance level, *** = 1% significance level.

Table 8 presents the results of a random effects regression with amount contributed as the dependent variable. The results indicate that after controlling for round, session, and player, the incentive type has no statistically significant impact on contribution.[17]

CONCLUSION

Our experimental tests of the hypothetical pay hypothesis indicate mixed results. Some results from the UG show that players do not play the same way in games involving hypothetical incentives as they do in games involving real incentives. Players in hypothetical sessions tended to offer less, that is, they played closer to the Nash equilibrium. This effect was not seen in the first round of play, where subjects' offers were statistically indistinguishable, but became more apparent in successive rounds of play. However, in APGG it appears that players generally respond the same way to hypothetical incentives as they do to real incentives (only one of our hypothetical sessions appeared statistically significantly different from the real session). These results are not significantly different from results found in other studies that focus on reducing the costs of economic experiments, in that certain types of experiments

can be played without paying substantial payoffs. While it does seem irrational to say that participants in economic experiments will always perform similarly with hypothetical and real incentives when asked, it is not that unrealistic to say that in some economic experiments, players will play with the same strategy. While the reasons for the differences are unclear, there are some possible explanations which can be drawn by looking at the differences in the two types of games.

First, there is a complexity factor to consider. Andreoni's game is undoubtedly more complicated to learn and play with a consistent strategy than the UG. Thus, a player may figure that if he is going to learn how to play APGG, they might as well play their best. Correspondingly, a game that is complex to learn may have a higher level of competitiveness, which could certainly be a driving factor in economic games. Thus, in APGG, players may be driven by an incentive to just "win" by getting a big payoff. Given the relative simplicity of the UG, playing for hypothetical incentives may appear to be rather mundane to participants, especially when the game is continued for ten consecutive rounds. There also does not appear to be as much of an incentive or drive to "win" in the UG. Thus, the amount of effort a player puts into an economic game may directly be reflected by how much they consider it to be a "game" in the traditional sense.

Another possibility has to do with the simultaneous play of the Andreoni game. Since players are making decisions at the same time, there is no rejection factor and no punishment factor in the Andreoni game. Thus players do not have to gauge their decisions according to possible responses of an unknown player, like the UG. In the UG, there is still a rejection and punishment factor in the hypothetical sessions. The only difference is that the rejection or punishment is no longer real, thus, players should be less concerned with being rejected.

The final and most significant aspect that should be considered is the idea of fairness in economic games. If it is concluded that using hypothetical incentives eliminates the self-interest aspect from the UG,[18] then either a fairness factor or fear of rejection factor should become the dominant strategy. If fairness is indeed an important factor in determining offers in the UG, then when real incentives are removed, there should be an expected decrease in the average offer. Since there is less of a self-interest factor, fairness should motivate players to move towards an equal split. However, the evidence presented here suggests that, if anything, removing real incentives causes the average offer to decrease, thus fear of rejection seems to be the more dominant factor. If there is no real money to be won, there is less of a concern of being rejected.[19] Similarly, if player 2 rejects unequal offers because they are not fair, then she would have less incentive to do this if the payoffs are hypothetical; however, our data suggests that offers are more likely to be accepted under real incentives after controlling for the amount offered.

The study presented here provides some preliminary evidence that hypothetical incentives in some economic experiments give accurate results. Further development of this idea could include several possibilities. First, to control for the possible "complexity" factor, a simpler version of a simultaneous type of game, such as the Prisoner's Dilemma, could be played with hypothetical and real incentives. Additionally, to further investigate fairness and competitiveness in experiments such as these, players could play these games with only certain members of the group being paid.

The evidence presented here suggests that in some economic experiments simply requesting that subjects behave in a manner they would if incentives were real may provide accurate results. Additional effort should be addressed at determining the factors that lead to situations in which hypothetical and real incentives are equivalent. With this knowledge, both the number of experiments that can be conducted, and the number of observations that can be recorded, could increase significantly.

APPENDICES

Instructions for Andreoni's Public Goods Game. Italicized items in brackets were inserted for the appropriate sessions.

Contribution Decision Game Instructions

On each of your direction sheets there is a number in the upper left-hand corner. This number denotes your player number. This will be your number for both experiments. It is imperative that you do not reveal this number to any other participant. (*This paragraph was the first paragraph in the instructions for the first game in every session, regardless of which game was being played*.)

The first [*second*] game that we will play involves dividing all the participants into randomly assigned groups of three. You will NOT know who the other members are in your group, and the group assignments will change each round. Please note that only the experimenters will know who the other two members in your group are.

Your group makes up what will be called a contribution group. The members of your group will each make individual contributions to a Group Investment. Your payoff from these contributions will depend on how much you invest and how much the other two members of your group invest.

It is your task to decide how many tokens you would like to put into the group investment. For each round, you can contribute any number of tokens you choose from 0 to 7. (Remember: the other two members of your randomly assigned group will be making similar decisions while you are making yours.)

If you would now turn over the next page, you will see the Payoff Matrix that will determine what your payoff will be (in cents) based on your contribution and the other group members contributions. You will not actually receive this payoff, but we ask that you play the game as if you were going to. [*Please note that in this game you will receive this actual payoff amount. At the end of the game we will total your payoffs from each round and pay this total (again, in cents).*] (Note that all members follow the exact same Payoff Matrix.) In each round you will use a contribution decision form to fill out the number of tokens you wish to contribute to the Group Investment. Please do not make any other markings on this form. The forms will then be collected by the experimenters, and each form will be randomly assigned to a group. The experimenters will then denote on the form what group number you were assigned for that round (this will change each round) and what the total contribution was by the other two members of the group. The contribution decision forms will then be returned to the players, at which point you should use the payoff matrix to determine what your payoff

for that round was, and mark it on the contribution decision form. The game will then continue for a pre-determined number of rounds known only to the experimenters.

If you will now turn over the next sheet on your desk you will find the record sheet for the first [*second*] game. After each round, please use this record sheet to fill out what group number you were in, what you contributed, what the total amount the other members of your group contributed, and what your payoff was, in the respective columns.

We will now provide an example to demonstrate how the Payoff Matrix works. Suppose that you decide to invest 2 tokens in the Group Investment. Also, suppose that one other player invests 4 and another player invests 6. Thus, your contribution is 2 and the total amount contributed by the other two members is 4+6=10. So, using the Payoff Matrix, going down the column marked 2 and across the row marked 10, your payoff would be 140. Similarly, the player who contributed 4 receives a payoff of 86 (4 down, 2+6=8 across), and the player who contributed 6 receives a payoff of 47 (6 down, 2+4=6 across).

Contribution Decision Form (Form used for Andreoni's Public Goods Game).

Contribution Decision Form

Player_____   Round_____

Remember: You can contribute anywhere from 0 to 7 tokens. Your payoff will be determined by the Payoff Matrix.

Please circle the number of tokens you would like to contribute in this round:

     0   1   2   3   4   5   6   7

My Payoff was_____

(To be filled out by experimenter)

Group Number_____

Other Players' Total Contributions_____

Payoff Matrix for Andreoni's Public Goods Game.

		Your Contribution							
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**
	0	0	1	3	6	9	10	11	10
	1	1	4	8	11	14	15	15	14
	2	5	9	14	18	20	21	20	17
	3	12	17	22	26	28	28	25	22
Total	**4**	21	28	33	36	37	35	32	27
Amount	**5**	34	40	45	48	47	44	39	32
Contributed	**6**	49	56	60	61	59	54	47	38
by Other	**7**	68	74	77	76	72	64	55	44
Two Group	**8**	90	95	96	93	86	76	64	51
Members	**9**	115	118	117	111	102	89	74	58
	10	143	144	140	131	119	103	85	66
	11	175	173	166	153	137	118	97	75
	12	210	205	193	177	157	134	109	84
	13	248	239	223	203	178	151	122	93
	14	290	276	256	230	201	169	136	103

Instructions for the Ultimatum Game. Italicized items in brackets were inserted for the appropriate sessions.

Proposition Decision Game Instructions

If you do not remember your player number it is again located in the upper left-hand corner of these instruction sheets. Once again, please do not reveal this number to any other participant. (*This paragraph was the first paragraph in the instructions for the second game in every session, regardless of which type of game was being played.*)

Based on your seating decision, you have been pre-assigned to be either a proposer or responder for the first [*second*] game that we will play. In this game, each proposer will decide how to divide up ten dollars between himself and a randomly matched responder. The randomly assigned responder will then be given the choice to accept the proposal, with each player receiving the amount that was proposed, or reject the proposal, with each player receiving nothing. The proposer can divide up the ten dollars however he chooses, down to the nearest nickel.

Please note that you will not actually be paid the amounts determined in this game, but we ask that you play the game as if you were. Remember if the offer is rejected by the responder, both the proposer and the responder receive nothing. [*At the end of this second game, one round will be randomly selected by a drawing. For this randomly selected round you will be paid the actual amount of cash that you received from that round. Remember if the offer is rejected by the responder, both the proposer and the responder will receive nothing for the randomly selected round.*]

Each round will begin with the proposers deciding how they want to divide the $10.00. This decision will be made on the proposition decision form. Please mark only

what you have decided to keep and what you have decided to offer. The experimenters will then collect the forms and randomly distribute them to the responders. Responders are then asked to fill out their player number and circle whether or not they have decided to accept the offer. If you will turn over the next page on your desk you will find a record sheet for the second game. If you are a responder, please fill out each round after you have made your decision. If you are a proposer, please fill out each round after the forms have been returned to you by one of the experimenters. The game will be repeated for a pre-determined number of rounds. Please note that although the responder will know the player number of the proposer, and the proposer will know the player number of the responder after the response, nothing will be known about the identity of either player.

Proposition Decision Form (Form used for the Ultimatum Game).

Proposition Decision Form

Proposer: Player_____   Round_____

Remember: You can divide $10.00 any way you choose, down to the nearest nickel. The amount you keep and the responder gets must add up to $10.00. If the responder accepts, the money will be split as suggested, if the responder rejects, both players will receive nothing.

Please denote the amount that you wish to keep for yourself and the amount that you propose giving to the responder:

I Keep_____

Responder Gets_____

(To be filled out by the responder)

Responder: Player_____

Responder Accepts? (Please Circle YES or NO)

YES / NO

NOTES

1. Note that a similar approach is used in this study in respect to UG, as only one randomly selected round is paid.
2. Bolle acknowledges that this method would not work for every type of economic game. He especially points out that games with high decision costs or ethical incentives are more likely to reduce a player's willingness to play if pay is going to be random. However, for simple games, such as UG, the method appears to have merit.
3. The Dictator Game is a variation of UG where whatever split is made by the proposer cannot be rejected by the responder.
4. Vickrey auctions are auctions that have the highest bidder pay the second highest bid. These auctions are believed to reveal a player's true willingness to pay.
5. See Appendix *Payoff Matrix for Andreoni's Public Goods Game*.
6. Based on the idea that a responder would receive 0 accepting or rejecting. In practice, this number is actually the minimal amount that a proposer can offer.
7. Fairness can also be interchanged with charity, the idea that the more the proposer offers, the better he feels. In theory, a person playing only based on this idea would offer the entire pot to the responder, which has been seen in practice. Regardless of the classification, fairness or charity, there is a negative effect on the Nash equilibrium.
8. The only guarantee of knowing exactly what your payoff will be without any risk of rejection is to offer everything.
9. The idea that a player would be concerned with fairness and fear of rejection seems to be unrealistic. If a proposer were playing "fair," the idea of being concerned with being rejected is an anomaly.
10. It was believed that asking participants to play a game with hypothetical incentives followed by one with real incentives would suffer from unwanted order effects.
11. For a detailed view of the instructions and forms for the Ultimatum Game see Appendices.
12. For a detailed view of instructions, forms, and the payoff matrix for Andreoni's game, see Appendices.
13. Recognizing that since subjects were randomly matched, each subject might affect the others, so that individual data may not be independent observations, we examine first round data from each session in addition to all of the data collected.
14. This is further confirmed by the results of Kolmogorov-Smirnov and Mann-Whitney tests which reject the hypothesis the equality of the distributions of offers between the real and hypothetical incentive sessions, when examining all rounds. When looking only at first round offers, you cannot reject the null hypothesis for these tests.
15. If it is accepted that offers are different given the different type of incentives, a means test comparing acceptance rates across sessions would be inaccurate, since acceptance is clearly affected by the amount offered.
16. The Kolmogorov-Smirnov and Mann-Whitney test confirm the statistical equality of the distributions.
17. At the suggestion of an anonymous reviewer, we conducted a simple Monte Carlo experiment to determine the statistical power of these hypothesis tests. Data was generated for simulated 'real' and 'hypothetical' sessions with the same variances as our actual data but different means. In order to consistently (greater than 95 percent of the time) reject the null hypothesis in RE regressions using this simulated data, the difference in the mean contribution needed to be greater than 0.6, while in our actual data the observed difference was 0.253. This implies that there may be insufficient experimental data to resolve a difference that may well exist.
18. A more likely conclusion is that it simply scales self-interest down.
19. It is unclear if this would actually cause a decrease in offers, however if players are still playing with any self-interest or level of competitiveness, it would be expected that offers should decrease.

REFERENCES

Andreoni, J. An Experimental Test of the Public-Goods Crowding-Out Hypothesis. *American Economic Review,* December 1993, 1317-1327.

Blumenschein, K., Johannesson, M., Blomquist, G., Liljas, B., and O'Conor, R. Hypothetical versus Real Payments in Vickrey Auctions, *Economics Letters,* October 1997, 177-180.

Bolle, F. High Reward Experiments without High Expenditure for the Experimenter? *Journal of Economic Psychology,* June 1990, 157-167.

Camerer, C. and Hogarth, R. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty,* December 1999, 7-42.

Camerer, C. and Thaler, R. Ultimatums, Dictators and Manners. *Journal of Economic Perspectives,* Spring 1995, 209-219.

Cameron, L. Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia. *Economic Inquiry,* January 1999, 47-59.

Forsythe, R., Horowitz, J., Savin, N., and Sefton, M. Fairness in Simple Bargaining Experiments. *Games and Economic Behavior,* May 1994, 347-369.

Harrison, G., Johnson, E., McInnes, M., and Rustrom, E. Risk Aversion and Incentive Effects: Comment. *American Economic Review,* June 2005, 897-901.

Holt, C., and Laury, S. Risk Aversion and Incentive Effects. *American Economic Review,* December 2002, 1644-1655.

_____. Risk Aversion and Incentive Effects: New Data without Order Effects. *American Economic Review,* June 2005, 902-912.

Kahneman, D. and Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica,* March 1979, 263-291.

Roth, A, Prasnikar, V., Okuno-Fujiwara, M., and Zamir, S. Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review,* December 1991, 1068-1095.

Sefton, M. Incentives in Simple Bargaining Games. *Journal of Economic Psychology,* June 1992, 263-276.

Slonim, R., and Roth, A. Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. *Econometrica*, May 1998, 569-596.

Thaler, R. The Psychology and Economics Conference Handbook: Comments on Simon, on Einhorn and Hogarth, and on Tversky and Kahneman. *Journal of Business,* October 1986, S279-S284.