

Pre-Play Contracting in the Prisoners' Dilemma

James Andreoni
Department of Economics
University of Wisconsin

and

Hal Varian
School of Information Management and Systems
University of California at Berkeley

May 1999
Current Version, July 8, 1999

Classification: Social Sciences, Economic Sciences

Corresponding Author: James Andreoni, Department of Economics, University of Wisconsin, Madison, Wisconsin 53706, andreoni@facstaff.wisc.edu

Manuscript Information: 16 Pages , 4 Figures, 0 Tables.

Abstract: We consider a modified Prisoners' Dilemma game in which each agent can offer to pay the other agent to cooperate. The subgame-perfect equilibrium of this two-stage game is Pareto efficient. We examine experimentally whether subjects actually manage to achieve this efficient outcome. We find an encouraging level of support for the mechanism, but also find some evidence that subjects' tastes for cooperation and equity may have significant interactions with the incentives provided by the mechanism.

1. Introduction

It is easy to write down simple games that have inefficient equilibria. The classic Prisoners' Dilemma is perhaps the simplest game with a dominant strategy equilibrium that is Pareto inefficient. Literally thousands of experiments on the Prisoners' Dilemma have been conducted across the social sciences.* While researchers often find some fraction of cooperation, the general result in these experiments is that the incentives to defect can be very powerful. When subjects are faced with a single shot of a Prisoners' Dilemma they seldom reach mutually cooperative outcomes.†

There is a growing number of studies that examine mechanisms for implementing Pareto efficient outcomes in games with inefficient Nash equilibria. Some of the first studies were conducted by Smith (7, 8) on the "Smith Auction," a mechanism which requires unanimity to implement "bids" from all players. While the theoretical properties of this mechanism are not well established, Smith nonetheless finds that his mechanism generates significant amounts of cooperation.‡ Other research into more well-understood mechanisms has generated less optimism. Atiyeh, Franciosi and Isaac (10) examined the dominant strategy Groves-Clarke mechanism, and were disappointed to find that fewer than ten percent of subjects reveal preferences truthfully. Others have considered Nash equilibrium mechanisms. Bagnoli and McKee (11) examined a "provision point" game in which private pledges are collected only if a threshold of total pledges is met. They find Pareto efficient outcomes are provided only 54 percent of the time. The results were even less encouraging when Bagnoli, Ben-David and McKee (12) turned to multi-unit provision point games. Harstad and Marrese (13, 14) examined the Groves-Ledyard mechanism. They found the Pareto efficient outcome was implemented less than twenty percent of the time. Chen and Plott (15) and Chen and Tang (16) also considered Groves-Ledyard mechanisms and found similarly disappointing results with a low "punishment parameter" in the mechanism. However, when they assigned a high punishment parameter the choices converged to the

*See Rapaport and Chammah (1) and Dawes (2) for reviews of these experiments in Sociology and Psychology. Roth (3) surveys some of the studies by economists.

†It is interesting to note, however, that subjects do respond to incentives to increase cooperation, despite the existence of a dominant strategy. Roth and Murningham (4), for instance, found that in a game with an uncertain end point the subjects were more cooperative the longer the anticipated horizon. Selten and Stoecker (5) and Andreoni and Miller (6) look at finitely repeated games and find that subjects will be more cooperative when they can build reputations.

‡Banks, Plott and Porter (9) generalized Smith's mechanism and find similar results.

Pareto efficient equilibrium in virtually every session.[§]

All of these prior experiments on mechanism design looked at dominant strategy or Nash equilibrium mechanisms.[¶] In this paper we take another approach and study a mechanism that implements the efficient outcome in a subgame perfect Nash equilibrium. Such subgame perfect mechanisms have gathered interest in the theoretical literature; see Moore and Repullo (20) for an extensive theoretical examination of implementability in subgame perfect equilibria. But as far as we are aware, this paper presents the first experiments on this type of mechanism.

The game we look at is modeled after work by Varian (21) regarding pre-play contracting. More precisely, we add another stage to the Prisoners' Dilemma game in which the players can make binding commitments to pay the other player some amount if he chooses to cooperate. Of course, it has long been known that the ability to make binding commitments eliminates the "dilemma" in the Prisoners' Dilemma. However, there has been surprisingly little written about the exact form that such binding commitments might take. A notable exception would be experimental studies of bargaining games. In these games subjects must agree on how to divide a fixed pie, and failure to agree results in the loss of the pie. When subjects alternate making binding offers, as suggested by Rubinstein (22), the sub-game perfect equilibrium involves the first player making an offer that is (barely) acceptable to the opponent. Hence, subgame perfection implies efficiency. As discussed in the review article by Roth (23), however, such subgame perfect behavior is seldom observed in these games.

In the experimental game we study here, we model the "commitment stage" explicitly in a two-stage game whose subgame-perfect equilibria are Pareto efficient. These equilibria imply a certain pattern of transfers between the agents that is reminiscent of a competitive equilibrium. In particular, in order to induce cooperation, each agent must be paid an amount at least as large as the amount he would receive if he were to defect. Roughly speaking, each agent is receiving his opportunity cost for cooperation.

We set a higher bar for ourselves by first "training" subjects on a basic Prisoners' Dilemma game, without a mechanism. This induces the normal amount of

[§]Chen and Tang (16) also consider the Walker mechanism, again with disappointing results. Chen (17) makes the conjecture that convergence under the high punishment parameter is due to the fact that only in this case is the mechanism supermodular. Supermodularity provides very robust stability properties that are consistent with learning dynamics. Cheng (18) explicitly considers the dynamic properties of the compensation mechanism, and finds that, with a slight modification, it too satisfies supermodularity.

[¶]Chen (18) provides an excellent review of this literature.

defection and frustration among subjects. Only then do we introduce the mechanism. We find this mechanism to be quite successful. By the end of the experiment about two-thirds of all subjects chose cooperation, which compares favorably with the experimental results for the other efficiency-inducing mechanisms described above. We see some systematic deviations from our predicted equilibrium, but this behavior is consistent with subjects having tastes for equity similar to those observed in other experiments. We conclude that the mechanism was successful at increasing efficiency, and that the subjects played reasonably close to the subgame perfect equilibrium.

2. Pay for Play

We begin by describing the exact form of our two-stage game, which we call Pay for Play. Imagine a standard Prisoners' Dilemma game with the strategy set {Cooperate, Defect}. We introduce a prior stage to this game where each agent can announce a nonnegative number, indicating the amount that he will pay the other agent if he chooses the Cooperate strategy in the second stage. This announcement is binding; once the agent offers a contract, he is obligated to carry it out.

Let us calculate the subgame perfect equilibrium of such a game. Consider the following example of an asymmetric Prisoners' Dilemma:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	6; 7	0; 11
	Defect	9; 0	3; 4

Now we add an “announcement stage” to this game where each agent simultaneously and independently announces how much he will pay the other agent if he cooperates. Player 1 announces a sidepayment of s_1 and Player 2 announces a sidepayment of s_2 . Given a pair of announcements the game becomes:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$6 - s_1 + s_2; 7 + s_1 - s_2$	$0 + s_2; 11 - s_2$
	Defect	$9 - s_1; 0 + s_1$	3; 4

If Player 1 cooperates, Player 2 receives 11 by defecting and 7 by cooperating. Therefore the minimal payment that would induce him to cooperate is 4. Similarly, the minimum payment to induce Player 1 to cooperate when Player 2

cooperates is 3. If these payments are announced, that is $(s_1, s_2) = (4, 3)$, then the second-stage game is transformed to:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	5; 8	3; 8
	Defect	5; 4	3; 4

Note that in this game it is a weakly dominant strategy for each player to choose Cooperate.

It can be shown $(s_1, s_2) = (4, 3)$, followed by mutual cooperation, is the unique subgame perfect equilibrium when the sidepayments can be any real number.^{||} In the experimental game that we consider, the sidepayments are restricted to be integers. This restriction adds a new subgame perfect equilibrium to the game, namely one where each agent pays 1 unit more than the break-even announcement. In the example we are considering here we would have $s_1 = 5$ and $s_2 = 4$ so the game becomes

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	5; 8	4; 7
	Defect	4; 5	3; 4

In this game it is strictly dominant strategy to Cooperate. This equilibrium is supported by the pessimistic expectations that if the other player is indifferent between cooperating and defecting, he will choose to defect. The other equilibrium is supported by the optimistic expectations that if the other player is indifferent between his two strategies he will choose to cooperate. Of course, it follows trivially that $(s_1, s_2) = (4, 4)$ or $(s_1, s_2) = (5, 3)$, along with mutual cooperation, are also equilibria.

The idea of adding a “contracting stage” to the Prisoners’ Dilemma is a variation on Varian’s (21) compensation mechanism. The idea is that each player offers to compensate the other for the “costs” that he incurs by making the efficient choice. Varian (21) shows that this sort of “compensation mechanism” is very powerful. It can be used to internalize nearly any sort of externality, resolve

^{||}See Varian (21) for a demonstration. Note that Ziss (24) showed that this is not true for all constellations of payoffs in the Prisoner’s Dilemma game. For the parameters used in this example, which are also those used in the experiment, the Pareto efficient allocation is the unique subgame perfect equilibrium.

public goods problems, regulate bilateral monopolies, etc. In addition to being robust, the mechanism yields outcomes that are competitive equilibria, with externalities priced at the appropriate efficiency price.

3. The Experiment

Our experiment, which was run on a computer network, is designed and presented to look like a card game. An example of the computer screen that the subjects saw is illustrated in Figure 1. Instructions used in the experiment are presented in the Appendix.

[FIGURE 1 ABOUT HERE]

Each game has two players with a pot of chips between them, as shown on the screen in Figure 1. Each chip is worth 5 cents. Player *A* has two playing cards, a “Push card” with a 6 on it, and a “Pull card” with a 4 on it. If player *A* plays his Push card, then 6 chips are pushed from the pot to the other player. If player *A* plays his Pull card, then he pulls 3 chips from the pot to himself. Player *B* has a push card of 7 and a pull card of 3 and has the same options. Players must move simultaneously. Note that this game has the same payoff structure as the Prisoners’ Dilemma game illustrated above—the dominant strategy is for each player to choose the Pull card, while the Pareto efficient strategy is for each player to choose the Push card.

To give the mechanism the greatest challenge, we first ran 15 rounds of this “Push-Pull” game with each subject randomly reassigned a different opponent each time. As is commonly observed, the players started out cooperating but soon switched to defecting. By the last round, most players were playing the Defect strategy of Pull.

We then switched to a new game called “Pay-for-Push”. In this game, each player independently names an amount of chips that he will pay the other player if the other player chooses the Push card. Each player manipulates a slider on a computer screen to indicate how many chips he would transfer to the other player if the other player chooses to Push. Once both players had committed to their payments, the amounts were revealed to the other player and we moved on to the Push-Pull game described above. We repeated this for 25 rounds. Hence, subjects played a total of 40 rounds, 15 of Push-Pull and 25 of Pay-for-Push.

We conducted 6 sessions in all, with 8 players per session, using a total of 48 subjects. Subjects 1–4 played against subjects 5–8. After each round, players

1–4 were reassigned to a different player in the 5–8 group. Each player had cards (4,6) or (3,7) and switched cards after each play of the game. Each session of the experiment was complete within an hour, and the subjects earned about \$9 to \$15 in the experiment, plus were paid \$3 for showing up.**

4. Results

Figure 2 shows the fraction of subjects choosing the to cooperate over all 40 iterations of the game. Look first at rounds 1 to 15, which are just the standard Prisoner’s Dilemma game. Over these rounds, 25.8 percent of moves are cooperative, which declines to 22.9 percent over rounds 10 to 15. This is consistent with results from other recent Prisoner’s Dilemma games.†† A surprising result from this part of the study, however, is that the propensity to cooperate differs significantly across the two types of players. Players in the role of *A* cooperate 19.2 percent of the time, while those in the Player *B* role cooperate 32.5 percent of the time. This difference is statistically significant ($t = 4.13$). The gap remains in the final five rounds of Prisoners’ Dilemma, with *As* at 16.7 percent and *Bs* at 29.2 percent cooperation ($t = 2.32$). What could explain this asymmetry? Player *As* have a push card of 6 while *Bs* have a push card of 7. One possibility is that the “warm glow” from giving may be higher for *Bs* since the gift to the other player is higher. At the same time, *As* have a Pull card of 4 while *Bs* have a Pull card of 3, so the cost of cooperating is also lower for the *Bs*. Hence, if there is an independent utility from the act of cooperating, as many have suggested, then we may expect more cooperation from *Bs*.

[FIGURE 2 ABOUT HERE]

Given this “training” at defecting, how do subject respond to the introduction of the Pay-for-Push mechanism? Figure 2 shows that the response is tentative at first, but after three rounds there is clear movement toward more cooperation. During the entire Pay-for-Push segment, 50.5 percent of all moves are cooperative,

**We recruited some subjects from a subject pool that had been used in a very different bargaining experiment several months earlier, but we do not know if any of our subjects had participated in an experiment before. Other subjects were recruited from economics principles courses. We were careful to recruit subjects before any mention had been made of Prisoners’ Dilemma in the classroom.

††For example, Roth and Murningham (4) find 10.1 percent cooperative moves, Cooper, De-Jong, Forsythe and Ross (25) find 20 percent, and Andreoni and Miller (6) find 18 percent.

and 54.5 percent are cooperative over the last five rounds of play. Again, however, there is a significant difference between *As* and *Bs*, with cooperation by *As* at 41.8 percent and by *Bs* at 59.2 percent ($t = 6.09$).

Clearly the mechanism is having some effect—cooperation has doubled. But the mechanism is also far from 100 percent successful. There are two reasons that this could be so. First, subjects could be failing to make the subgame perfect sidepayments in first phase of mechanism, and second, when they receive the appropriate sidepayment they fail to respond optimally. We examine both of these possibilities next.

4.1. Sidepayments — Phase 1 of the Mechanism

Figure 3 shows the average sidepayment made during the experiment for both *As* and *Bs*. The subgame perfect prediction is that Player *A* should be offering 3 or 4, and Player *B* should be offering 4 or 5. We see that actual behavior deviates some from this prediction, especially for *Bs*. Over the 25 rounds, Player *As* average offer is 3.39, which rises to 3.57 for the final five rounds, which is consistent with the prediction. For *Bs*, however, the overall average sidepayment offered is 3.30 on average, rising to 3.52 over the last five rounds, but is well short of the predicted 4 to 5.

[FIGURE 3 ABOUT HERE]

It is not obvious why sidepayments would differ in this way. One possibility is that this could be another manifestation of pure tastes for cooperation we saw in the regular Prisoners' Dilemma game. Alternatively, some subjects could have been using the sidepayments to try to even out earnings in the mutually cooperative outcome. This could be accomplished if Player *A*'s sidepayment is one greater than Player *B*'s. In the subgame perfect equilibrium, however, *B*'s sidepayment is one greater than *A*'s. It could be that some Player *As* could have tried to increase their sidepayment, and some *Bs* decrease theirs, in the hopes that tastes for cooperation or equity would be enough to enforce the mutually cooperative outcome, leading to the choices observed.

Let's examine this by looking more carefully at the offers made. First, ask how often an offer at or above the equilibrium offer was made. Overall, 63.5 percent of all sidepayments were at or above the predicted amounts, rising to 84.9 percent for the last 5 rounds. This fraction, however, is higher for the Player *As* than for Player *Bs*— 78.7 percent for *As* and 48.3 percent for *Bs*, which is a significant

difference ($t = 8.91$). Over the last five rounds, however, the fraction is similar and the difference is insignificant — 85.8 percent for *As* and 84.2 percent for *Bs* ($t = 0.36$). How often did a player make an offer strictly above the predicted amount, that is, above 3 for *As* and above 4 for *Bs*? Overall this happened 39.9 percent of the time, but again it is far more likely to happen among the *As*. The difference is 57.5 percent for *As* and 22.3 for *Bs*, which is significant ($t = 13.32$), and it grows to 65.8 versus 12.5 percent for the last five rounds ($t = 10.10$). Hence, by the end of the experiment players in both roles are equally likely to make an offer that should generate a cooperative reply, but the surplus in the offer is significantly higher in offers made by Player *As*. This indeed indicates a bias in play toward more “fair” allocations.

4.2. Cooperation — Phase 2 of the Mechanism

Next we ask how subjects respond to the sidepayments offered in phase 1. Figure 4 shows how often subjects cooperated when they were offered a sidepayment at least as high as the subgame perfect Nash equilibrium prediction. The average over the 25 rounds of the mechanism is 67.9 percent. Hence, when it is optimal to cooperate, two thirds of subjects do. What about when the offer makes cooperation a dominant strategy? Now the fraction who cooperate when given such an offer rises to 77.4 percent.

[FIGURE 4 ABOUT HERE]

Is there a difference between the Player *As* and Player *Bs* in this regard? Disaggregating this we find that *As* respond optimally to good offers 63.1 percent of the time, while *Bs* do 71.0 percent. For offers generating a dominant strategy, *As* cooperate 71.6 percent of the time and *Bs* 79.7 percent. Again, there is a difference in the actions of players in the two roles, albeit a smaller difference than in the offers.

What about the propensity to defect when the sidepayments offered are so low that defection is the best reply to either of the other player’s choices? Given such a bad offer, 79.9 percent of people will defect. Note this is only slightly higher than the rate of defection in the ordinary Prisoners Dilemma game, which as we reported earlier was 77.1 percent.

Again, let’s separate the behavior of the Player *As* and *Bs*. When *As* got a bad offer, that is of 3 or less, they defected 78.1 percent of the time. When *Bs* get a bad offer, that is of 2 or less, they defected 84.4 percent of the time. In contrast

to the earlier results, these differences are not significant ($t = 1.59$). Hence, any significant difference between *As* and *Bs* tends to vanish when they fail to get a good offer in the mechanism.

4.3. Summary — A Closer Look at Equity

We have now seen that, while the mechanism works reasonably well, there is significant asymmetry between the play of *As* and *Bs*. In Prisoners' Dilemma without the mechanism, *Bs* are significantly more cooperative than *As*. With the mechanism, *As* make sidepayments that have a significantly higher surplus. How do we interpret the asymmetry?

To understand the asymmetry of behavior, we should begin by examining the asymmetry in the payoff matrix in the standard Prisoners' Dilemma game. When each player is choosing between his Push card and Pull card, he is weighing whether to be nice or be selfish. Player *As* can give 6 to the other player at a cost of 4 to themselves. This implies a price of cooperation of $2/3$. Player *Bs* face a similar decision, but they can give 7 at a cost of 3 implying a price of $3/7$. Since $3/7 < 2/3$, we would predict *Bs* to be more cooperative, which is what we see.

How would this effect translate to the Pay-for-Push game? Implicit in the last paragraph is that individuals have a natural desire to be nice in these situations— a hypothesis that has received a fair degree of support in prior experiments.^{‡‡} If there is a utility component in cooperation above and beyond the monetary gain, then it may be that one need not pay the full “opportunity cost” of cooperation in order to elicit cooperation in the mechanism. Since *Bs* face a lower price than *As*, they are already more inclined to cooperate; hence a sidepayment below the subgame perfect payments is more likely to induce cooperation for the *Bs* than for the *As*. This means that, relative to the subgame perfect sidepayment, *As* should get higher payments than *Bs*. Unfortunately, this is exactly the opposite of what happened.

This means that simple demands for being nice will not be enough to explain this asymmetry. A more complete explanation can be found by looking at final payoffs. If the subgame perfect Nash equilibrium is reached, the payoffs will be fairly unequal — *As* will earn 8 and *Bs* 5. Another finding from experiments

^{‡‡}See footnote 1 for references in the broad social science literature. Within economics see, for example, Palfrey and Rosenthal (26) for a discussion of social dilemmas, Andreoni and Miller (6) on Prisoners' Dilemma, Andreoni (27) and Palfrey and Prisbrey (28) on public goods, and Andreoni and Miller (29) for general tastes for giving.

on games with unequal equilibrium payoffs is that subjects tend to dislike unequal payoffs, especially when on the losing end (see (30), (31)). If subjects do indeed dislike inequality, then this should put pressure on sidepayments to equal out earnings, not simply to pay the opportunity cost of cooperation. This in turn should suppress sidepayments by B and increase sidepayments by A . This is exactly what was observed in the experiment.

It is fair to conclude, therefore, that the kinds of taste for equity seen in other experiments may be at work here and, as in other settings, is interacting with the incentives of the mechanism. The asymmetry of the payoffs seems particularly powerful in affecting the choices of our subjects.

5. Conclusion

We considered a subgame perfect Nash equilibrium mechanism that was applied to the Prisoners' Dilemma. This is perhaps the simplest form of this mechanism that could be implemented in an experiment. We saw that, overall, the mechanism was largely successful at implementing Pareto efficient allocations. We found that players made offers of sidepayments that should induce cooperation about 63.5 percent of the time. When such offers were received, subjects responded with cooperation nearly 70 percent of the time. Accounting for the usual levels of noise and confusion in experiments, we find these results encouraging for this mechanism.

The presence of concerns for equity were also found in our data, and they appeared to interact with the incentives presented in the mechanism. Especially important was the apparent effort of subjects to use the sidepayments to “undo” the inequality implied by the asymmetric payoff matrix, as well as to cover the opportunity costs of cooperation. A symmetric game, while less revealing about the ability of subjects to learn the mechanism, may eliminate some of the concerns with inequality that entered our experiment. Alternatively, one could follow the experiences of Chen and Plott (15) and use a steeper payoff space with a greater gain from cooperation in hopes that this could swamp the concerns for equity and get an even stronger result for the mechanism. These could be interesting topics for future research.

Appendix: Subjects' Instructions

Instructions for Zenda

Zenda is a simple card game that you play with one other person. Each game of Zenda will consist of 15-25 rounds of play. Players are matched up randomly each round, so that you will play a different person each round. All of your choices and earnings in the experiment will be confidential.

In Zenda, you will be playing for chips. The value of the chips corresponds to cash earnings for you. In particular, each chip you earn is worth 5 cents. So if you earn 5 chips in a round, you earn 25 cents in the round. If you earn 10 chips, you earn 50 cents in that round. The earnings that you make each round will be totaled by the experimenter and paid to you privately and in cash at the end of the experiment. No other subject will know your earnings.

After each round is finished, a dialog box will be displayed informing you of this fact and asking you to wait until the other players are finished. Please click on the "OK" button as soon as you have read and understood the material since the system will wait until everyone has clicked "OK" before it proceeds. Dialog boxes will be displayed at other times during the game; after you have read and understood a dialog box, click "OK".

Please do not talk to any other player or look at any other player's screen. If you have a problem, please raise your hand and someone will come to help you. We expect the experiment to last about 50 minutes.

Information about you

The first thing that you see will be a panel that asks for information about you. The University requires that we collect this information since we are going to give you money. This information is not recorded as part of the experimental data; it is only there to satisfy University rules about dispersing money. In the experiment you are only identified by a "player number" and we maintain strict anonymity. No one in the experiment will ever know your name, your choices, or your earnings.

Phase 1 of Zenda

There are two phases to Zenda. Phase 1 is called Push-Pull. When you play Push-Pull, you will see two cards in front of you, two cards in front of the other player, and a pile of chips between the two of you. The pile of chips between you is the "pot"; it is the source of the payments.

Your cards are labeled Push and Pull. You can choose to play a card by clicking on it with the mouse. When you choose a card it will be highlighted but your choice will not be final until you click the “Confirm choice” button. If you choose the pull card, then you will pull the number of chips on that card from the pot to your pile of chips. If you choose the push card, then you will push the number of chips on that card from the pot to the other player.

Note that the chips that you push or pull come from of the pot in the middle of the table, not from either player’s pile of chips. When both players have made their decision to push or to pull, you will see a message appear telling you what has happened and your earnings will be displayed. When all players have made their choices, you will see a panel and a beep will announce the start of a new round. Click on the “OK” button on the panel to start playing the new round.

An example

Suppose that your push card is a 6 and your pull card is a 4. Then if you choose to push, the other player will get 6 chips from the pot in the middle of the table. If you choose to pull, then you will get 4 chips from the pot.

Suppose that the other person has a push card of 7 and a pull card of 3. Then if he pushes, you will get 7 chips from the pot. If he pulls, you will get no chips from the pot.

The total number of chips that you end up with depends on the choices made by you and the choices made by the other player.

Summary Phase 1 of Zenda

Choose to push or pull and click “Confirm choice.” If you push, the other player gets the number of chips on your push card; if you pull, you get the number of chips on your pull card. When you see a dialog box that tells you the round has ended, click on “OK” so that the play can proceed. You will play 15 rounds of Phase 1.

Phase 2 of Zenda

In this phase you have a new option that we call Pay-for-Push. Everything else about Zenda will be the same as in Phase 1. In particular, during each round of Phase 2 you will be randomly paired with another subject each time you play Zenda.

On your screen you will see a new button labeled “Confirm payment” and a slider. You can use the slider to offer a payment to the other player to encourage

them to choose to play the push card.

You set this payment by moving the slider up or down. As you do this you will move chips from your pile of chips to a pile in front of the other player. When you are satisfied with your decision about how much you are willing to pay the other player, you click on your “Confirm payment” button. Neither player will see how much the other player is willing to pay until both players have clicked on their “Confirm payment” button. Once you’ve seen how much the other player has offered to pay you to push, you can decide whether to push or to pull. Your payoff will be as before but now the payment you make will be subtracted from your pile of chips if the other player chooses to push. If the other player chooses to pull, then you will get your payment back. Likewise, if you choose to push, the payment offered to you by the other player will be added to your earnings, and subtracted from the other players earnings.

Summary of Phase 2 of Zenda

In each round of play you will move your slider to determine how many chips you want to pay the other player to Push. Once you have decided this, you click on your “Set Payment” button. When both players have clicked their “Confirm payment” button, each will be able to see how much the other player has offered. At that point, each player can choose whether to push or pull as before. When both players have clicked their “Confirm payment” button, each player sees the payoffs and a new round begins.

You will play 25 rounds of Pay-for-Push.

Things to remember

1. You play against a new person every time.
2. You should click “OK” as soon as you have read and understood a message.
3. You are playing for real money; each chip is worth 5 cents.

Acknowledgements: We are grateful to the National Science Foundation for financial support. We also thank Yan Chen for helpful comments.

References

1. Rapaport, A., & Chammah, A. M. (1965), *Prisoner's Dilemma*, Ann Arbor: University of Michigan Press.
2. Dawes, R. M. (1980), "Social Dilemmas," *Annual Review of Psychology* **31**, 169–93.
3. Roth, A. E. (1988), "Laboratory Experimentation in Economics: A Methodological Overview" *The Economic Journal* **98**, 974–1031.
4. Roth, A. E., & Murnighan, J.K. (1978), "Equilibrium Behavior and Repeated Play of the Prisoner's Dilemma," *Journal of Mathematical Psychology* **17**, 189–198.
5. Selten, R., & Stoecker, R. (1986), "End Behavior in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach," *Journal of Economic Behavior and Organization* **7**, 47–70.
6. Andreoni, J., & Miller, J. H. (1993), "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *Economic Journal* **104**, 570–585.
7. Smith, V. L. (1979), "An Experimental Comparison of Three Public Goods Decision Mechanisms," *Scandinavian Journal of Economics* **81**, 198–215.
8. Smith, V. L. (1980), "Experiments with a Decentralized Mechanism for Public Good Decisions," *American Economic Review* **70**, 584–599.
9. Banks, J. S., Plott, C.R., & Porter, D. (1988), "An Experimental Analysis of Unanimity in Public Goods Provision Mechanisms," *Review of Economic Studies* **60**, 301–322.
10. Attiyeh, G., Franciosi, R., & Isaac, R. M. (1999), "Experiments with the Pivot Process for Providing Public Goods," *Public Choice*, forthcoming.
11. Bagnoli, M., & McKee, M. (1991), "Voluntary Contribution Games: Efficient Provision of Public Goods," *Economic Inquiry* **29**, 351–66.
12. Bagnoli, M., Ben-David, S., & McKee, M. (1992), "Voluntary Provision of Public Goods: The Multiple Unit Case," *Journal of Public Economics* **47**, 85–106.

13. Harstad, R. M., & Marrese, M. (1981), "Implementation of Mechanism by Processes: Public Good Allocation Experiments," *Journal of Economic Behavior and Organization* **2**, 129–151.
14. Harstad, R. M., & Marrese, M. (1982), "Behavioral Explanations of Efficient Public Good Allocations," *Journal of Public Economics* **19**, 367–383.
15. Chen, Y., & Plott, C. R. (1996), "The Groves-Ledyard Mechanism: An Experimental Study of Institutional Design," *Journal of Public Economics* **59**, 335–364.
16. Chen, Y. & Tang, F. (1998), "Learning and Incentive-Compatible Mechanisms for Public Goods Provision: An Experimental Study," *Journal of Political Economy* **106**, 633–662.
17. Chen, Y. (1997), "Supermodularity of Nash-efficient Public Goods Mechanisms." working paper, University of Michigan.
18. Cheng, J.Q.(1998), "Essays on Designing Economic Mechanisms," Ph.D Dissertation, University of Michigan.
19. Chen, Y. (1998), "Incentive-Compatible Mechanisms for Pure Public Goods: A Survey of Experimental Research," working paper, University of Michigan, 1998, forthcoming in C. Plott and V. Smith, eds., *Handbook of Experimental Economics Results*.
20. Moore, J. & Repullo, R. (1988), "Subgame Perfect Implementation," *Econometrica* **47**, 1191–1220.
21. Varian, H. R. (1994), "A Solution to the Problem of Public Goods When Agents are Well-Informed," *American Economic Review* **84**, 1278–1293.
22. Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica* **50**, 97–109.
23. Roth, A.E. (1995), "Bargaining Experiments," in J.H. Kagel and A.E. Roth, eds., *The Handbook of Experimental Economics*, Princeton: Princeton University Press.
24. Ziss, S. (1997), "A Solution to the Problem of Public Goods When Agents are Well-Informed: Comment," *American Economic Review* **87**, 231–235.

25. Cooper, R., DeJong, D.V., Forsythe R., & Ross, T.W. (1996), "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games," *Games and Economic Behavior* **12**, 187-218.
26. Palfrey, T.R., & Rosenthal, H. (1988), "Private Incentives in Social Dilemmas: The Effects of Incomplete Information and Altruism," *Journal of Public Economics* **35**, 309-32.
27. Andreoni, J.(1995), "Cooperation in Public Goods Experiments: Kindness or Confusion?" *American Economic Review* **85**, 891-904.
28. Palfrey, T.R., & Prisbrey, J.E. (1997), "Anomalous Behavior in Public Goods Experiments: How Much and Why?," *American Economic Review* **87**, 829-846.
29. Andreoni, J., & Miller, J.H. (1998), "Giving According to GARP: An Experimental Test of the Rationality of Altruism," working paper, University of Wisconsin.
30. Prasnikar, V., & Roth, A.E. (1992), "Considerations of Fairness and Strategy: Experimental Data From Sequential Games," *Quarterly Journal of Economics*, 865-88.
31. Andreoni, J., Brown, P.M., & Vesterlund, L. (1999), "What Makes an Allocation Fair? Some Experimental Evidence," working paper, University of Wisconsin.

Figure Legends:

Figure 1: Sample of the Computer Screen Seen by Subjects

Figure 2: Cooperation by Rounds. Rounds 1–15 Prisoners' Dilemma, Rounds 16–40 with Mechanism.

Figure 3: Average Sidepayments by Round

Figure 4: Probability of Cooperation when Receiving a Good Offer

Illustrations:

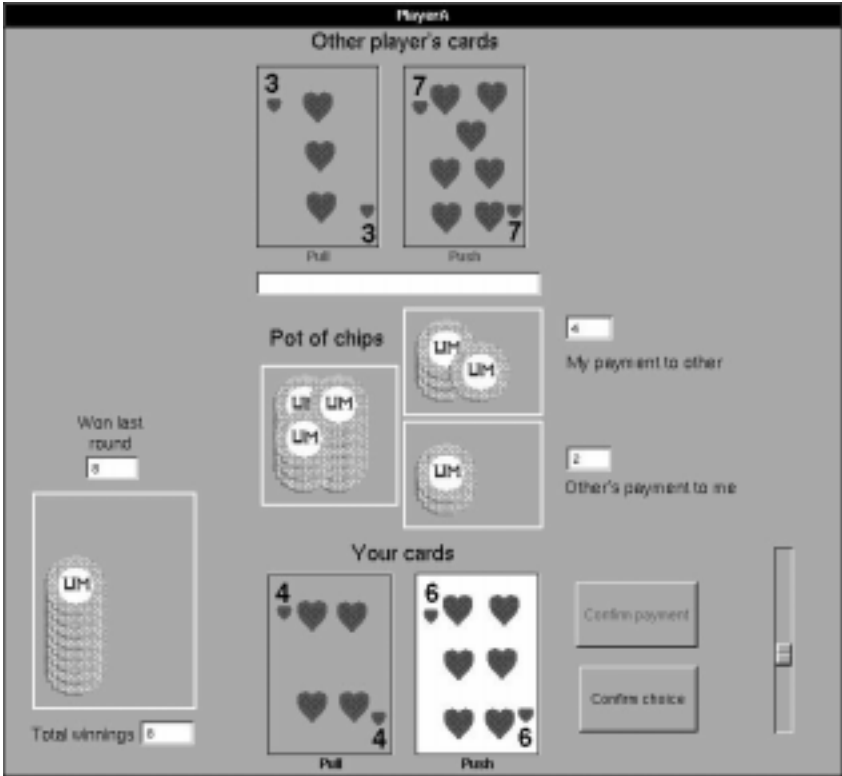


FIGURE 1: Sample of the Computer Screen Seen by Subjects

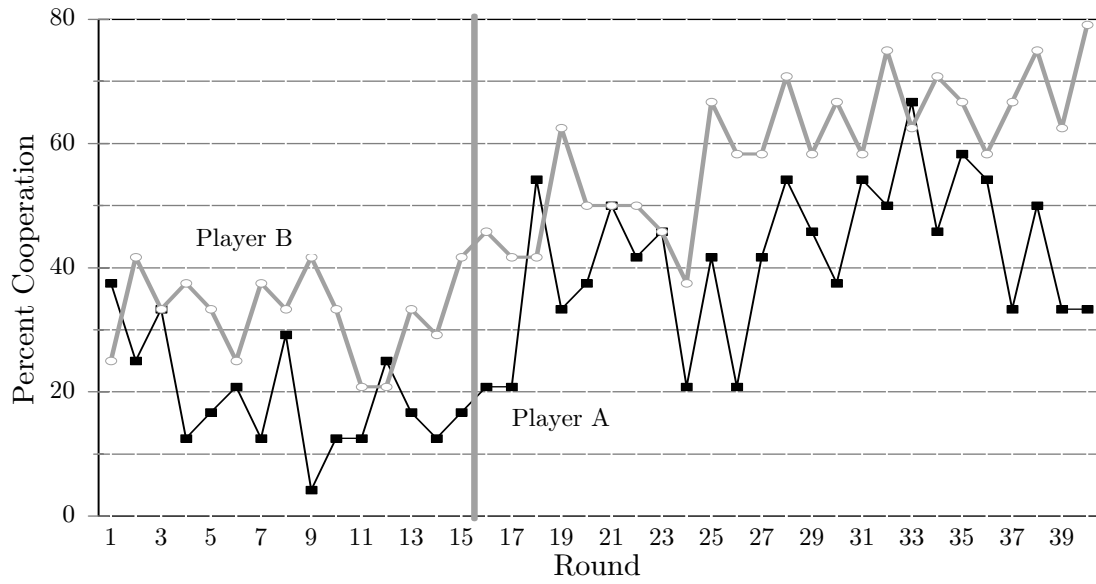


FIGURE 2: Cooperation by Rounds.
 Rounds 1–15 Prisoners' Dilemma, Rounds 16–40 with mechanism.

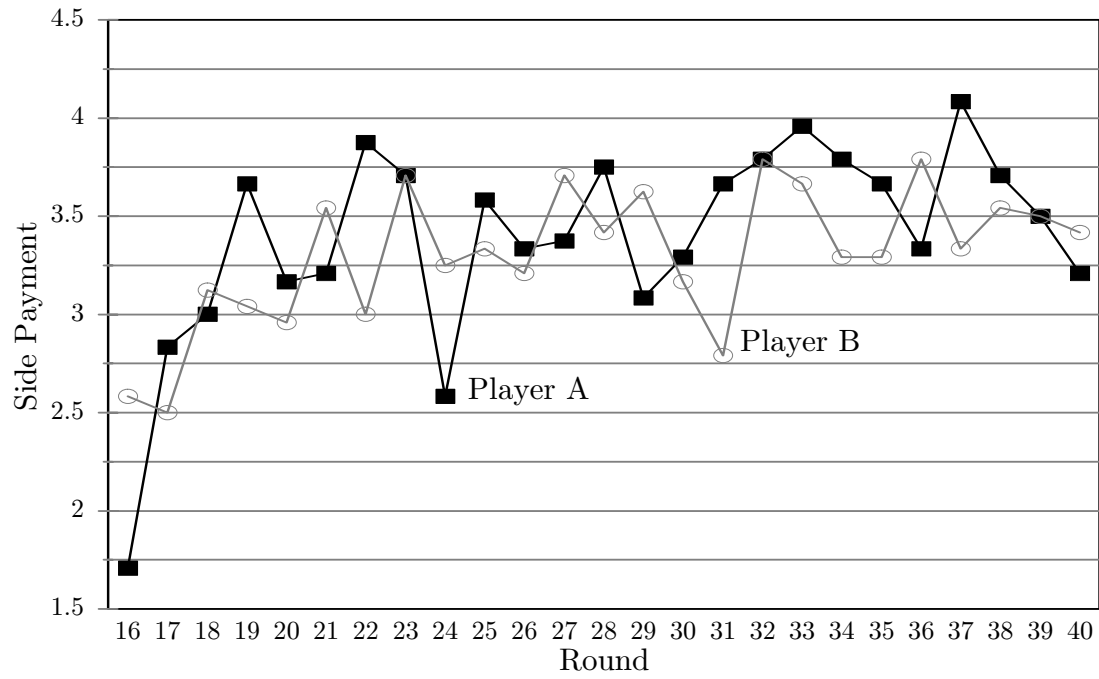


FIGURE 3: Average Sidepayments by Round

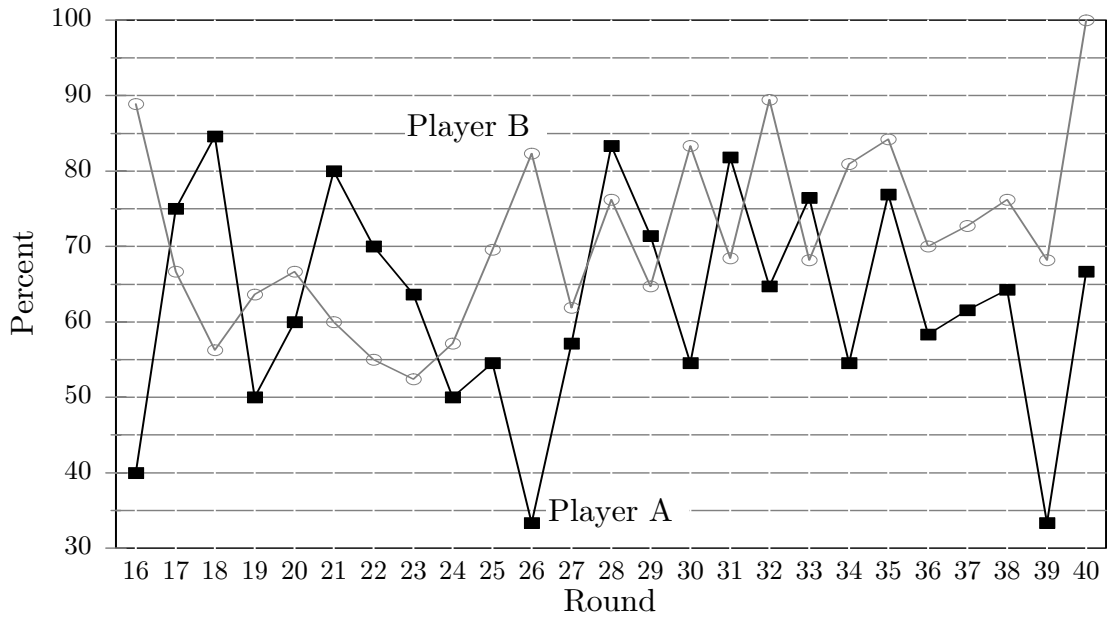


FIGURE 4: Probability of Cooperation when Receiving a Good Offer