

NBER TECHNICAL WORKING PAPER SERIES

DETERMINISTIC AND STOCHASTIC PRISONER'S DILEMMA GAMES:  
EXPERIMENTS IN INTERDEPENDENT SECURITY

Howard Kunreuther  
Gabriel Silvasi  
Eric T. Bradlow  
Dylan Small

Technical Working Paper 341  
<http://www.nber.org/papers/t0341>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
August 2007

We appreciate helpful discussions in designing the experiments and comments on earlier drafts of this paper by Colin Camerer, Vince Crawford, Rachel Croson, Robyn Dawes, Aureo DePaula, Geoff Heal, Charles Holt, David Krantz, Jack Ochs, Al Roth and Christian Schade. We also benefited from helpful comments from participants at the Workshop on Interdependent Security at the University of Pennsylvania (May 31-June 1 2006) and at the Workshop on Innovation and Coordination at Humboldt University (December 18-20, 2006). We thank George Abraham and Usman Hassan for their help in organizing the data from the experiments. Support from NSF Grant CMS-0527598 and the Wharton Risk Management and Decision Processes Center is gratefully acknowledged. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Howard Kunreuther, Gabriel Silvasi, Eric T. Bradlow, and Dylan Small. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Deterministic and Stochastic Prisoner's Dilemma Games: Experiments in Interdependent Security  
Howard Kunreuther, Gabriel Silvasi, Eric T. Bradlow, and Dylan Small  
NBER Technical Working Paper No. 341  
August 2007  
JEL No. C11,C12,C22,C23,C73,C91

**ABSTRACT**

This paper examines experiments on interdependent security prisoner's dilemma games with repeated play. By utilizing a Bayesian hierarchical model, we examine how subjects make investment decisions as a function of their previous experience and their treatment condition. Our main findings are that individuals have differing underlying propensities to invest that vary across time, are affected by both the stochastic nature of the game and even more so by an individual's ability to learn about his or her counterpart's choices. Implications for individual decisions and the likely play of a person's counterpart are discussed in detail.

Howard Kunreuther  
Operations and Information Management  
The Wharton School  
University of Pennsylvania  
3730 Walnut Street, 500 JMHH  
Philadelphia, PA 19104-6366  
and NBER  
kunreuther@wharton.upenn.edu

Gabriel Silvasi  
The Wharton School  
University of Pennsylvania  
3730 Walnut Street, 500 JHMM  
Philadelphia, PA 19104-6366  
silvasig@wharton.upenn.edu

Eric T. Bradlow  
The Wharton School  
University of Pennsylvania  
3730 Walnut Street, 761 JMHH  
Philadelphia, PA 19104-6366  
ebradlow@wharton.upenn.edu

Dylan Small  
The Wharton School  
University of Pennsylvania  
3730 Walnut Street, 464 JMHH  
Philadelphia, PA 19104-6366  
dsmall@wharton.upenn.edu

## 1. Introduction

Interdependent security (IDS) games are social multi-player games with stochastic payoffs where each player must decide whether or not to mitigate his or her own risks. More specifically, each player knows that even if she fully protects herself by investing in a risk-reducing measure, she may still be subject to indirect losses by being contaminated by one of the other players who chose not to invest in similar measures. As Heal and Kunreuther (2005) have shown, there are a wide variety of problems that fit into the IDS framework ranging from investing in airline or port security, protecting oneself against disease through vaccinations, individual users incurring software security costs when connected to a network, and divisions of firms undertaking risk-reducing measures to avoid a catastrophic loss that could cause the entire firm to become insolvent or bankrupt.

IDS problems can have multiple Nash equilibria or just a single Nash equilibrium depending on the tradeoffs between the marginal costs associated with investing in the protective measure and the marginal benefits from reducing direct and indirect losses (Kunreuther and Heal 2003). This paper presents the results of controlled laboratory experiments involving two players in an IDS game where there is a single Nash equilibrium. If both players invest in a risk-reducing measure, there is no chance that either will suffer a loss. However, if either or both players do not invest in the protective measure, then there is some likelihood (hence the stochastic part) that both individuals will suffer a loss. Furthermore the dominant strategy for both players, if they are risk neutral, is *not* to invest in the protective measure even though had they taken this step, their expected values would have been higher than if they had not. In this sense our IDS game can be viewed as a stochastic prisoner's dilemma (SPD) game.

The prisoner's dilemma problem is a paradigmatic example highlighting a broader class of problems that fall under the heading of social dilemmas. In such situations each individual receives a higher payoff for making a socially defecting choice (e.g. polluting the environment) than for pursuing a socially cooperative choice no matter what the other individuals in society do. Furthermore all individuals are better off if they all cooperate than if they all defect (Dawes 1980). In his classic piece on the tragedy of the commons, Hardin (1968) showed that the failure of each individual to restrain themselves from utilizing a resource could lead to its depletion (e.g., a fish population or a forest from which timber is harvested) or global changes (e.g., climate change due to excessive greenhouse gas emissions). Interest in the tragedy of the commons dilemmas has grown steadily in the last fifty years

because social scientists have pointed out that many of the major problems facing society fall under this rubric.<sup>1</sup>

Our interest in this research area is both substantive and methodological. From a substantive perspective, we would like to compare the behavior of individuals in a standard deterministic prisoner's dilemma (DPD) *supergame* of ten periods with their behavior in an SPD supergame based on an IDS model. Is there more or less cooperation (investment in protection) in an SPD game versus a DPD game and what are the factors influencing the level of cooperation? We are also interested in how feedback affects cooperation in a stochastic game in which each player is provided with information on the action of one's counterpart and the payoffs for both players. Specifically, we compare behavior in this full feedback stochastic game with behavior in a stochastic game where there is feedback on the payoffs for both players but not on the action of one's counterpart. In this partial feedback game, a player may or may not be able to infer her counterpart's action based on the payoffs received by both players.

From a methodological perspective, we utilize a Bayesian hierarchical model coupled with marginal analyses to understand more fully the factors that influence individuals to invest in protective measures over time and whether there are significant differences in behavior when payoffs are deterministic versus stochastic. The use of a hierarchical model is important to properly account for the correlation among repeated decisions of the same player. To the best of our knowledge this is the first time that Bayesian hierarchical methods have been applied to the analysis of repeated prisoner's dilemma game experiments.

The remainder of the paper proceeds as follows. The next section provides a brief summary of theoretical and empirical studies on prisoner's and social dilemmas with a focus on recent experiments in a noisy or stochastic environment. We then describe in Section 3 a general formulation for an interdependent security (IDS) two-player game, showing under what conditions it takes the form of a stochastic prisoner's dilemma game. After characterizing our experimental design for a set of controlled laboratory experiments (Section 4), we then specify a set of between treatment hypotheses and test them using regression analysis (Section 5). In Section 6, we build a Bayesian hierarchical model that enables us to test hypotheses with respect to within-subject behavior. Section 7 summarizes the findings, discusses their prescriptive implications and suggests areas for future research.

---

<sup>1</sup> For an insightful set of recent research articles on social and commons dilemmas in the context of resource problems see Ostrom et al (2002).

## 2. Background on Deterministic and Stochastic Prisoner's Dilemma Games

Here we lay out a basic taxonomy that provides a general research prospectus for repeated, multi-player stochastic prisoner's dilemma games. In particular, we point out where previous research falls in this structure, what holes are remaining, and which ones are satisfied by this research.

To anticipate what will be discussed in the next subsections, we focus on multi-period games and extend this framework to address whether the outcomes are deterministic versus stochastic. This issue is not as simple as it might appear because there are many aspects of a supergame that might or might not be stochastic. We then address whether subjects have the ability to learn about their counterparts' choices.

### *Deterministic Multi-period Games and the Emergence of Cooperation*

We first consider deterministic prisoner's dilemma (DPD) games. In a single period DPD game, defecting is the only Nash equilibrium and experiments have shown that players *learn* to play this Nash equilibrium in a series of games in which a player is matched with a different player in each period (e.g., Cooper et al., 1996). But in supergames, in which a player is matched with the same player for repeated periods, players learn to reciprocate cooperative behavior as they gain experience (Selten and Stoecker, 1986; Andreoni and Miller, 1993; and Hauk and Nagel, 2001). Cooperation tends to break down near the end of the supergame however. In a tournament setting, Axelrod and Hamilton (1981) and Axelrod (1984) showed that the tit-for-tat (TFT) strategy generates a fair amount of cooperation and outperforms other strategies.<sup>2</sup>

Kreps *et al.* (1982) have addressed the theoretical issue of whether it can be rational for players to cooperate in a finitely repeated game. They show that information asymmetries can yield a significant amount of cooperation in a rational equilibrium. Specifically, if each player perceives that there is some probability that the other player would like to cooperate for whatever reason and will play a strategy such as TFT that punishes her counterpart for not cooperating, then it will be optimal for both players to cooperate.

---

<sup>2</sup> A tit for tat strategy is one where a player cooperates on the first move and does whatever the other player does on every move thereafter.

### ***Stochastic versus Deterministic Games***

In a paper closely related to our study, Bereby-Meyer and Roth (henceforth B&R) (2006) compared cooperation in a multi-period DPD to cooperation in a multi-period stochastic prisoner's dilemma game in which the payoffs were random, but the expected payoffs were that of the DPD and the players learned the action that their counterparts took (i.e., there was full feedback). Subjects played 20 ten-period supergames and were informed of their own and their counterparts' payoffs and actions after each period. In the probabilistic conditions, the payoffs were determined by a binary lottery with a probability  $p$  of winning \$1 and probability  $(1-p)$  of winning \$0 with the value of  $p$  for each player depending on whether or not she and her partner cooperated. The expected values in the probabilistic condition were set equal to those in the deterministic condition.<sup>3</sup> B&R hypothesized that because the SPD provides only partial reinforcement for cooperating when one's counterpart cooperates, players would learn to cooperate more slowly (if at all) in the SPD compared to the DPD. B&R confirmed their hypothesis and showed that players' decisions in the SPD at time  $t$  are affected by the lottery that determined their random payoff at time  $t-1$  in addition to their counterpart's action at time  $t-1$ .

### ***Partial Feedback versus Full Feedback***

In the usual DPD and B&R's SPD, players know what actions their counterparts have taken in previous rounds. But in many real situations, decision makers are uncertain about their counterparts' actions. Motivated by this issue, Bendor *et al.* (1991) conducted a multi-period prisoner's dilemma game tournament modeled after Axelrod (1984) but with random payoffs and where players only learned about their own payoff. Players received no feedback regarding either their counterpart's action or payoff. Bendor *et al.* found that the TFT strategy, which outperformed other strategies in Axelrod's DPD tournament, fared rather poorly in this SPD tournament with partial feedback. Axelrod and Dion (1988) note that when there is uncertainty in outcomes, then cooperation may avoid unnecessary conflict but can invite exploitation. Axelrod (1984) and Donniger (1986) also presented results of tournaments of SPDs with partial feedback, finding that TFT could still perform well if there is only a small amount of noise in the payoffs.

---

<sup>3</sup> As a control case B&R designed a "deterministic plus sunspots" game in which payoffs were deterministic but players learned the outcome of two binary lotteries that did not influence the outcomes but had the same distributions as lotteries in the probabilistic condition.

Bendor (1987, 1993), Molander (1985), and Mueller (1987) studied SPDs with partial feedback from a theoretical perspective. Bendor (1993) showed that, although uncertainty about one's counterpart's actions hinders cooperation in some circumstances, there are other situations in which the uncertainty can enhance cooperation by allowing reciprocating but untrusting strategic players to begin cooperating because of unintended consequences.

### ***The Contributions of Our Experiments***

Our experiments are the first that we know of to compare side-by-side a DPD multi-period game, an SPD with full feedback (SPD-FF) multi-period game and an SPD with partial feedback (SPD-PF) multi-period game. Our SPD game with full feedback differs from B&R's in two important respects. First, in our SPD, individuals face the possibility of a loss rather than a gain, and are provided with a given amount at the start of each supergame. Second, in our SPD, the loss to both players is determined by the same lottery, whereas in B&R's SPD, the two players' outcomes are determined by independent lotteries. The shared losses in our SPD, which are a feature of IDS games, may affect players' psychological reaction to the outcome of the game and the other player's action. In our SPD game with partial feedback, players learn only their counterpart's *loss* and not their counterpart's *action*. For some combinations of a player's strategy and the losses of both players, it is *possible to infer* what actions one's counterpart took; for other combinations, one is clueless. To our knowledge, our experiments are the first to consider an SPD with partial feedback game among live players rather than a tournament with pre-entered strategies. Furthermore, for the SPD-FF and SPD-PF games, we vary the probability of experiencing a loss so that we can decompose the direct effects of the existence of stochasticity from the magnitude of the likelihood of the event.

Our first goal is to compare the overall levels of cooperation (investment in protection) in the information conditions DPD, SPD-FF and SPD-PF. We study this in Section 6. Our second goal is to understand the source of any differences in cooperation between the three information conditions in terms of how players respond to different situations. We study this in Section 7. First, we describe the IDS paradigm, our experimental design and hypotheses.

### **3. IDS Games**

To motivate our experiments in the context of interdependent security models we focus on an example of two identical individuals,  $A_1$  and  $A_2$ , each maximizing her own expected value in a one-

period model and having to choose whether to invest in a protective measure. Such an investment by individual  $i$  costs  $c$  and reduces the probability of experiencing a *direct loss* to 0. Let  $p$  be the probability of a direct loss to an individual who does not invest in protection. If an individual experiences a direct loss, there is some probability  $q \leq p$  that the other individual will also experience an indirect loss from this individual even if she has invested in protection. For example, an apartment owner who has invested in a sprinkler system to prevent fire damage may still suffer a loss indirectly from a neighboring unit that does not invest in this form of protection and experiences a fire. The direct or indirect loss to each player is  $L$ .<sup>4</sup>

Let  $Y$  be the assets of each individual before she incurs any expenditures for protection or suffers any losses during the period. Assume that each individual has two choices: invest in protection, **I**, or do not invest, **NI**, and makes her decision so as to maximize expected value. The four possible expected outcomes from these decisions are depicted in Table 1:

**[INSERT TABLE 1 HERE]**

To illustrate the nature of the expected returns consider the upper left hand box where both individuals invest in security (**I, I**). Then each individual incurs a cost of  $c$  and faces no possible losses so that each of their net returns is  $Y-c$ . If  $A_1$  invests and  $A_2$  does not, then this outcome is captured in the upper right hand box (**I, NI**). Here  $A_1$  incurs an investment cost of  $c$  but there is still a chance  $q$  that a loss will occur to  $A_2$  so that  $A_1$ 's expected loss from damage from a negative externality is  $qL$ . The lower left box (**NI, I**) has payoffs which are just the mirror image of these.

Suppose that neither individual invests in protection (**NI, NI**) – the lower right hand box in Table 1. Then each has an expected return of  $Y - pL - (1-p)qL$ . The expected losses can be characterized in the following manner. The term  $pL$  reflects the expected cost of a direct loss. The second term reflects the expected cost from an indirect loss originating from the other individual ( $qL$ ) and is multiplied by  $(1-p)$  to reflect the assumption that a loss can only occur once. In other words, the risk of contamination only matters to an individual when that individual does not experience a direct loss.

Assuming each individual wants to maximize her expected returns, the conditions for her to invest in protection are that  $c < pL$  and  $c < p(1-q)L$ . The first constraint is exactly what one would expect if the individual could not be contaminated by the other person. Adding a second individual tightens the

---

<sup>4</sup> We recognize that a more general formulation would allow for  $p$ ,  $q$ ,  $L$ , and  $Y$  to vary across players. However, the general taxonomy and theory is not restricted by our simplifications. Empirical findings under more general settings are an important area for future research.



constraint by reflecting the possibility of contamination should this person decide not to invest in protection. The resulting Nash equilibrium (NE) for this IDS model can be determined as follows:

- If  $c < p(1-q)L$  then **(I, I)** is a NE
- If  $c > pL$  then **(NI, NI)** is an NE
- If  $p(1-q)L < c < pL$  then both **(I, I)** and **(NI, NI)** are NE

An IDS game becomes an SPD game when  $pL + (1-p)qL > c > pL$  so that **(NI, NI)** is a dominant solution but both individuals would be better off if they had decided to invest in protection **(I, I)**. For the experiments described below, we set  $q = 1$  so that if one individual suffers a loss, the other individual is certain to also experience this same loss. We also choose values of  $p$  and  $L$  so the IDS game takes the form of an SPD game.

#### 4. Experimental Design

The two-person experiments were carried out in the behavioral laboratory of a large, northeastern university using a web-based computer program. The pool of subjects recruited for the experiment consisted primarily of undergraduate students, though a small percentage of subjects were graduate students and students from other area colleges. A more detailed demographic breakdown of the subject pool is given in Appendix A. The studies were run with three different experimental conditions containing between three and seven pairs of subjects participating in specific sessions. A session consisted of a set of supergames, each consisting of 10 periods. The computer program randomly paired the subjects before the start of each supergame.

A person played a 10-period supergame with his/her anonymous partner; and at the conclusion of the supergame the person was then told that she would be randomly paired before the start of the next supergame. The number of supergames in each session ranged from three to ten depending on how long the session ran and how rapidly the pairs of players were able to complete each supergame.<sup>5</sup> More than half the participants participated in exactly eight supergames.

Each subject was given an initial surplus of 300 “talers” (described below) at the beginning of every supergame. Before the experiment began, every subject was told that each supergame consisted of 10 periods. The number of supergames was not announced at the beginning of the experiment nor was the final supergame announced when it began. Subjects were also told that at the end of the entire

---

<sup>5</sup> It was necessary for all pairs to complete a 10-period game before one moved to the next game.

session, one supergame and one pair playing that supergame would be chosen at random, and each individual from the selected pair would receive the dollar equivalent of his/her final payoff from that 10 period game. The lucky pair received these payments in addition to the fixed fee of \$8-\$12 (depending on the length of the session) that each person received for participating in the experiment.

The initial surplus and payoffs presented to the subjects during the experiment were in an artificial currency called “talers,” and these were converted to dollars (10 talers=\$1) at the end of the experiment for the randomly selected lucky pair. The average earnings per person from the game for the pairs chosen at random to receive their final payoff from a supergame was \$25.55. Screen shots of the instruction pages for all conditions, as well as decision and payoff screens, are presented in Appendix B. We next describe each of these experimental treatment conditions in detail.

***Information Condition 1: Deterministic Prisoner’s Dilemma (DPD) Game***

Subjects in the DPD condition were presented the payoff matrix depicted in Appendix B (Figure B1). In this condition, the cost of investing is  $c=12$  talers. A loss of  $L=10$  talers (in addition to any investment costs) occurs for both players if exactly one player does not invest. A loss of  $L=16$  talers occurs for both players if both players do not invest. Both individuals would be better off if they had each invested rather than not invested; however (NI, NI) is the Nash equilibrium.

***Information Condition 2: Stochastic Prisoner’s Dilemma Game with Full Feedback on Counterpart’s Decision (SPD-FF)***

The stochastic conditions replicated the effect of  $p$  (the probability of a negative random event each period as described in Table 1) by means of a random number. In order to understand the impact of the probability and magnitude of a loss on behavior in the SPD game, we ran sessions where  $p=.2$ ,  $p=.4$  and  $p=.6$ . The cost of investing in protection remained at  $c=12$  for each of these cases; however the respective losses were set at  $L= 50, 25$  and  $19$  so that the expected payoff matrices were essentially equivalent across the three experimental conditions.<sup>6</sup> This was done so that we could isolate the impact of the change in loss probability on one’s decision to invest and not have it be confounded by the magnitude of the expected loss.

---

<sup>6</sup> For  $p=.2$  the expected losses were: **I,I** $=(-12,-12)$  **NI, I** $=(-10, -22)$  **NI, NI** $=(-18, -18)$ . For  $p=.4$  the expected losses were: **I,I** $=(-12,-12)$  **NI, I** $=(-10, -22)$  **NI, NI** $=(-16, -16)$ . For  $p=.6$  the expected losses were: **I,I** $=(-12,-12)$  **NI, I** $=(-11, -23)$  **NI, NI** $=(-16, -16)$ .

The exact payoff matrix presented to subjects in the SPD-FF condition, sub-treatment  $p=0.2$ , is depicted in Appendix B (Figure B5); the payoff matrix for sub-treatment  $p=0.4$  is depicted in (Figure B8); and the payoff matrix for sub-treatment  $p=0.6$  is depicted in (Figure B9).

At the end of each period  $t$  each player was told whether his/her counterpart had chosen **I** or **NI** (hence full feedback). If either (or both) players had chosen **NI**, then the computer would draw a random number from 1 to 100 and highlight this number on a table on each subject's screen indicating whether or not a loss had occurred. Each player was then shown his/her cumulative balance in talers and his/her counterpart's cumulative balance for period  $t$  and all previous periods in the supergame (See the example in Appendix B).

***Information Condition 3: Stochastic Prisoner's Dilemma Game with Partial Feedback on Counterpart's Decision (SPD-PF)***

This game is identical to the SPD-FF game except that after each period  $t$  the random number generates an outcome even if both players invested in protection (i.e. **(I,I)**). Each player is then told whether or not she suffered a loss but not what action her counterpart had taken. Each player learns only her own cumulative balance in talers for period  $t$  and all the previous periods. For some outcomes in period  $t$  it is *possible* for a player to determine what action her counterpart had taken that period. For example, if Player 1 invests in protection (**I**) and suffers a loss, then she knows that Player 2 must not have invested (**NI**). In other cases it is impossible for the other player to learn what her counterpart has done. For example, if a green (no-loss) random number is drawn, then there is no loss whether or not either player invested in protection. Table 2 summarizes the twelve combinations of investment-color configurations that could arise each period, and whether Player 1 can learn whether or not her counterpart has invested in that period (a symmetric table exists for Player 2).

**[INSERT TABLE 2 HERE]**

***General Overview of the Data Collected for All Conditions***

Table 3 indicates the number of individuals who participated in experiments in each of the three treatment conditions and as a function of  $p$  for the SPD experiments.

**[INSERT TABLE 3 HERE]**

Although each individual played several supergames, every supergame played by a particular individual was within the same treatment condition. (e.g. SPD-FF,  $p=.4$ ). Thus for each subject there exists a vector of person-level covariate data about that subject (age, gender, race, etc.), plus a vector of

treatment-level covariate data for the subject (deterministic condition vs. stochastic full-feedback vs. stochastic partial-feedback, and  $p=0.2$  vs.  $p=0.4$  vs.  $p=0.6$ ), plus a series of 10-period supergame vectors, each of which contain the data about (a) the decision the subject made in each period (**I** or **NI**), (b) the decision her counterpart made in each period (**I** or **NI**), (c) the color that appeared on the random-number grid for that round (red, orange, or green), and (d) the number of points (talers) deducted from the subject's account each period. In addition to these four decision-level pieces of information collected during the experiment, we also added a binary variable indicating whether or not the subject was able to learn the decision of her counterpart; this "learn" indicator can be calculated based on Table 2.

## **5. Analyses of Between Treatment Hypotheses**

We first postulate between treatment hypotheses that can be tested with marginal analyses and simple regressions. These include (a) how levels of investment differ between the DPD, SPD-FF and SPD-PF conditions and (b) how levels of investment differ as the probability of a negative random event ( $p$ ) increases from 0.2 to 0.4 to 0.6.

### **5.1 Specific Between-Treatment Hypotheses**

*H1: The probability of investment will be greater in the DPD game than in either of the SPD games.*

Bereby-Meyer and Roth (2006), (B&R) found evidence that there was less cooperation in an SPD-FF game than in a DPD game. They ascribe this finding to the fact that stochastic games only provide partial reinforcement for cooperation. We expect to find a similar set of results.

*H2: The probability of investment will be greater in the SPD-FF game than in the SPD-PF game.*

We hypothesize that the lack of complete information in the SPD-PF game and hence the inability to learn what one's counterpart has done in many scenarios (see Table 2) will limit the emergence of stable cooperation between the two players.

*H3: For either SPD-FF or SPD-PF, the probability of investment will be greater in the  $p=0.4$  treatment than in the  $p=0.2$  treatment, and will be greater in the  $p=0.6$  treatment than in the  $p=0.4$  treatment.*

As  $p$  increases, the likelihood that a non-investing subject will experience a loss increases. There is considerable evidence from other studies that experiencing a loss increases the incentive to invest in protection. Kunreuther (2006) has shown that homeowners are likely to purchase

earthquake or flood insurance after a recent disaster even when they indicate that the probability has not increased (flood) or may even be lower in the immediate future (earthquake). Because subjects who do not invest in the previous period are more likely to experience a loss as  $p$  increases, this would lead them to invest in the next period, other things being equal.

*H4: In situations in which a subject's counterpart invests and this can be learned by the subject, there is a higher proportion of investment in the next round in the SPD-FF game than the SPD-PF game.*

In the SPD-FF game, a subject is explicitly told what his or her counterpart's investment decision in the previous round was. On the other hand, in the SPD-PF game, even when a subject can learn what his or her counterpart's investment decision was, he or she needs to infer the decision rather than being explicitly told. We hypothesize that the explicit learning of the SPD-FF compared to the implicit learning of the SPD-PF will make subjects more likely to reciprocate if their counterparts cooperate. As shown in Table 2 a subject can only learn that his or her counterpart has invested in an SPD-PF game when the subject has not experienced a loss (Scenarios 1 and 8).

## **5.2 Marginal Analyses**

Before formally testing the between treatment hypotheses, we first explore the data at its marginal level. Table 4 describes the proportion of times individuals invested in protection (cooperated) in the different conditions of our experiment. Figure 1 displays these mean investment levels along with the .05 and .95 quantiles of the distribution of subject investment levels in a condition, where the subject investment level for a given subject is the proportion of times the subject invested across the subject's supergames. At a glance, the investment proportion tended to be highest in the DPD, with roughly similar investment proportions in SPD-FF and SPD-PF. In the SPD conditions, the investment proportion increases with the probability of loss. There is a substantial amount of heterogeneity of investment among different individuals within a given condition; for example for DPD, the .05 quantile of investment proportion is 0.10 and the .95 quantile of investment proportion is .95. Although there is substantial heterogeneity among individuals within a condition, there are clear patterns of different mean investment levels across conditions.

**[INSERT TABLE 4 AND FIGURE 1 HERE]**

We now explore how the probability of investing changes as the number of the supergame varies from 1 to 8 (few individuals played more than 8 supergames) for each of the information conditions. These analyses mimic many of those in B&R. Figure 2 shows these probabilities for periods 1, 5, 8 and 10 (beginning, middle, near-end, and end rounds).

**[INSERT FIGURE 2 HERE]**

For periods 1, 5 and 8, the probability of investing in the DPD goes up gradually as players gain experience with the supergames. On the other hand, for the SPD-FF and the SPD-PF, the probability of investing goes *down* slightly as the players gain experience with the supergames. This is a stronger effect of uncertainty hindering the evolution of cooperation than B&R found. In their experiments, cooperation increased more slowly in an SPD than a DPD but still increased in the SPD. For period 10, the probability of investing is similar across all conditions. Cooperation declines in the last period compared to other periods in the DPD game, presumably because there is no incentive to cooperate in the last period.

We next explore how the probabilities of investment change as the period increases from 1 to 10. For brevity, we show these probabilities in Figure 3 for the first and last supergames – 1 and 8.

**[INSERT FIGURE 3 HERE]**

For the DPD in supergame 1, investment generally declines gradually as the periods increase. For the DPD in supergame 8, investment declines gradually from periods 1 to 8 and then declines sharply in periods 9 and 10. This behavior of declining investment as a function of period mirrors the findings of Selten and Stoecker (1986), Andreoni and Miller (1993), Hauk and Nagel (2001) and B&R (2006). For the SPDs in supergame 1, investment declines from period 1 to 3 and then stays relatively flat. For the SPDs in supergame 8, investment generally declines gradually. This is similar to B&R's finding that in an SPD, investment declines gradually as the period increases. In other words there is less of a drop in investment from period 1 to period 10 for the SPD than the DPD game.

### **5.3 Regression Analyses**

To more formally test the between treatment hypotheses H1-H4, we fit a regression of the proportion of times<sup>7</sup> each subject invested as a function of the subject's information condition (DPD, SPD-FF or SPD-PF), the subject's probability of loss condition if in one of the stochastic information

---

<sup>7</sup> We recognize that taking  $\text{logit}(\text{proportion})$  would yield a dependent variable more in line with the assumption of ordinary regression, however, we wanted to provide exploratory results on the original scale. Model-based analyses, presented at the end of Section 5 and in Section 6, are based on a logistic regression, more appropriate to the 0/1 nature of the data.

conditions SPD-FF or SPD-PF (( $p=.2$ )\*stochastic, ( $p=.4$ )\*stochastic and ( $p=.6$ )\*stochastic, where stochastic = 1 if in a stochastic condition and 0 otherwise) and interactions between the information condition and the probability of loss condition. The interaction terms were not significant ( $p$ -value for F-test = 0.63) and hence were dropped from the regression. Table 5 shows confidence intervals for a variety of interesting contrasts in the proportion of investment between conditions.

**[INSERT TABLE 5 HERE]**

The data provide strong evidence for H1 – there is strong evidence (in Table 5, R1-R6) that there was a higher mean investment proportion in the deterministic information condition than in either of the two stochastic information conditions for each of the three probability of loss levels. There is substantially more investment in the DPD in some cases. For example, the estimated difference is 0.29 between the DPD and (SPD-FF,  $p=.4$ ) ( Table 5 R3) , and 0.25 between the DPD and (SPD-PF,  $p=.4$ ), (Table 5 R4).

The data do not support H2 – there was slightly *less* investment in the full feedback SPD than the partial feedback SPD, the estimated difference is 0.04 [95% credibility interval: (0.00, 0.09)]; the sign of the difference between full feedback and partial feedback SPD is the opposite of what we hypothesized in H2, albeit the effect is quite small.

We now consider hypothesis H3 that subjects’ mean investment level increases as the probability of loss increases. Comparing the probability of loss conditions among subjects playing a given SPD game, there was substantially more investment when the probability of loss was .6 than when it was .4 or .2; the estimated difference in mean investment proportion is 0.16 for  $p=.6$  compared to  $p=.2$  (Table 5 R8) and 0.12 for  $p=.6$  compared to  $p=.4$  (Table 5 R9), and both differences are statistically significant ( $p<0.05$ ). There was slightly more investment for  $p=.4$  compared to  $p=.2$  – the estimated difference is .05, but the difference is not statistically significant. Thus, there is evidence for part of H3 that the subjects’ mean investment level increases as the probability of loss increases from  $p=.2$  or  $p=.4$  to  $p=.6$ , but there is no strong evidence that the mean investment level increases as  $p$  increases from .2 to .4.

We now consider hypothesis H4, i.e. that the explicit learning of the SPD-FF makes subjects more likely to reciprocate an investment by their counterpart than the implicit learning of the SPD-PF. Table 6 shows the probabilities of investment in the SPD-FF compared to the SPD-PF when the counterpart’s previous decision can be learned, broken down by the subject’s previous decision and the counterpart’s previous decision. The first two rows of the table (Table 6 R1, R2) support hypothesis H4

in that there is a higher proportion of investment in the SPD-FF than in the SPD-PF when the counterpart's previous decision to invest is learned.

To gauge the statistical significance of these differences, we need to account for the fact that (1) the periods and supergames at which subjects face a situation such as both players having invested in the previous round might differ in their distribution between the SPD-FF and the SPD-PF and (2) subjects contribute repeated investment decisions and these repeated investment decisions are likely to be correlated due to heterogeneity of investment propensity between subjects (see Figure 1 and Section 7 for evidence of this). To account for issue (1), we conducted a logistic regression analysis, regressing each investment decision (1 if invests, 0 if does not) as a dependent variable on dummy variables which encode the previous decision by the subject, the previous decision by the counterpart and whether the previous decision could be learned; interactions of these dummy variables with the subject's information condition; the subject's probability of loss condition; dummy variables for the periods; dummy variables for the supergame; interactions between the information condition and the period; interactions between the information condition and the supergame number; interactions between the probability of loss condition and the period; and interactions between the probability of loss conditions and the supergame. To account for issue (2) that subjects' repeated observations are likely to be correlated, we use the generalized estimating equations (GEE) method of Liang and Zeger (1986) with a working independence hypothesis. The inferences from a GEE analysis are asymptotically valid regardless of the true correlation among subjects' repeated observations.

There are two ways that H4 can be examined: when both the subject and counterpart previously invested and the counterpart's previous investment decision could be learned (H4A) and when the subject did not invest in the previous period but the counterpart did and the counterpart's previous investment decision could be learned (H4B). To test H4A for a given period and supergame, we want to determine whether the log odds ratios for investing in SPD-FF compared to SPD-PF is greater than 0. More specifically we want to determine whether a linear combination of these log odds ratios that puts equal weight on each period/supergame combination is greater than 0. The *p-value* for this test is <0.0001 and a 95% confidence interval for this linear combination of odds ratios is (0.48, 0.88); thus, there is strong support for hypothesis H4A. Similarly, to test H4B for a given period and supergame, we want to determine whether the log odds ratio for investing in SPD-FF compared to SPD-PF is greater than 0. The *p-value* for the test that is analogous to our test of H4A is 0.006 and a 95% confidence



interval for the linear combination of odds ratios is (0.10, 0.64); thus, there is strong support for hypothesis H4B.

As mentioned in the discussion of hypothesis H4, the distinction between the different ways of learning in the SPD-FF versus the SPD-PF when the counterpart chooses not to invest, is confounded by the fact that the subject always experiences a loss when he or she learns that the counterpart did not invest in the SPD-PF. Nevertheless, it is of interest to compare SPD-FF and SPD-PF for these situations. When one could learn that the counterpart chose not to invest in the previous period and when the subject also chose not to invest in the previous period, then there was a higher level of investment in the SPD-PF than the SPD-FF (Table 5 R4) and this was statistically significant at the 0.05 level in the GEE analysis. But when the subject chose to invest in the previous round, there was a higher level of investment in the SPD-FF than the SPD-PF analysis (Table 5 R3), although this difference was not statistically significant at the 0.05 level.

We note that applying a similar GEE analysis to that used to test H4 to H1-H3 produces the same conclusions as above; there is strong evidence for H1 and part of H3 but not for H2.

#### **[INSERT TABLE 6 HERE]**

In summary, we found that investment levels were significantly lower in both the SPDs than the DPD. Investment levels were comparable in the full feedback and partial feedback SPD. Investment levels for the SPDs increased as the probability of loss increased from .2 to .4 to .6, with the sharpest increase occurring from .4 to .6. Subjects were more likely to reciprocate investments by their counterparts in the SPD-FF when subjects explicitly learn their counterpart's investment decision than in the SPD-PF, even when subjects implicitly learned that their counterparts invested.

## **6. Bayesian Analyses of Within Treatment Hypotheses**

In order to gain further insight into the reasons for the differences in investment across the treatment conditions, we formulated a set of within treatment hypotheses that could be tested using Bayesian analyses.

### ***6.1 Specific Hypotheses***

*H5: There will be substantial variability between subjects in their tendency to invest, holding other factors fixed.*

This hypothesis characterizes how subjects within the same experimental treatment differ between each other in their aggregate investment levels, holding other factors fixed. It is supported by the findings of Andreoni and Miller (1993) who showed that some individuals were naturally cooperative even if they knew their counterpart was non-cooperative while others were non-cooperative even if they were aware their counterpart was cooperative.

The next set of hypotheses examine within-subject comparisons on how individuals change their investment levels over time with regards to (a) their decision to invest or not invest in the previous period, (b) their counterpart's decision to invest or not invest in the previous period (when this can be learned), and (c) the interaction between investment decisions in the previous period and whether or not an individual experienced a loss in the previous period.

*H6: In all three information conditions, subjects will tend to maintain the same investment behavior over time, holding their counterparts' decisions fixed.*

There is a large body of empirical and experimental evidence suggesting that individuals maintain the status quo even though they may be able to improve their expected profits by modifying their behavior (Samuelson and Zeckhauser, 1988; Kahneman, Knetsch and Thaler 1991).

*H7: In all three information conditions, a subject will be more likely to invest if he/she learns that his/her counterpart invested in the previous period.*

In spite of people's tendency to persist in their actions (as stated in H6), there is evidence that people can learn to cooperate in repeated prisoner's dilemma games (Axelrod, 1984). A theoretical model as to how cooperation can emerge in repeated prisoner's dilemma games was presented by Kreps et al (1982). Such cooperation is expected to emerge here.

*H8: If a subject experienced a loss in the previous period, he or she will be more likely to invest in the current period than if he or she had not experienced a loss in the previous period, holding all other conditions fixed.*

We can further divide H8 into four (interaction) sub-hypotheses.

*H8A: In the SPD-FF, when a subject invested in the previous period and his or her counterpart did not invest in the previous period, then the subject will be more likely to invest in the current period if he or she experienced a loss in the previous period than if he or she did not experience a loss.*

*H8B: In the SPD-FF, when a subject did not invest in the previous period and his or her counterpart invested in the previous period, then the subject will be more likely to invest in the current period if he or she experienced a loss in the previous period than if he or she did not experience a loss.*

*H8C: In the SPD-FF, when a subject did not invest in the previous period and his or her counterpart also did not invest in the previous period, then the subject will be more likely to invest in the current period if he or she experienced a loss in the previous period than if he or she did not experience a loss.*

*H8D: In the SPD-PF, when a subject did not invest and could not learn his or her counterpart's investment decision in the previous period, then the subject is more likely to invest if he or she experienced a loss than if he or she did not experience a loss.*

The reasoning behind Hypothesis H8 was discussed in the context of H3. There is considerable evidence from other studies that experiencing a loss increases the incentive to invest in protection. Note that the situation described in H8D is the only situation in the SPD-PF when a subject's response to a loss holding all other conditions fixed can be studied (see Table 2).<sup>8</sup>

*H9: If a subject did not invest in the previous period, experiencing a loss (compared to not experiencing a loss) will have a bigger effect on the odds of a subject investing if the subject's counterpart invested in the previous period compared to if the subject's counterpart did not invest in the previous period.*

We hypothesize that a subject experiencing a loss is likely to feel more regret if she did not invest in the previous period and this was the sole cause of the loss compared to the situation where she did not invest in the previous period but her counterpart was partially responsible for the loss. Note that H9 can only be tested in the SPD-FF.

---

<sup>8</sup> In Table 2 Scenarios 7 and 10 are the two cases where a subject has not invested in the previous period, experiences a loss and cannot learn what her counterpart has done. These two scenarios can be compared with Scenarios 9 and 12 where the individual has not invested, not suffered a loss and has no idea what her counterpart has done.

## 6.2 A Bayesian Hierarchical Model for Individual Investment Decisions

To examine these hypotheses we build a Bayesian hierarchical model for how subjects make investment decisions as a function of their previous experience and the treatment condition they are in.<sup>9</sup> The following notation characterizes a model of individual (investment) choice:

Let  $i = 1, \dots, I$  index subjects participating in the experiment ( $I = 936$ );

$t = 1, \dots, T$  index periods in a supergame ( $T=10$ );

$r = 1, \dots, R$  index the round of the supergame played in one participant session (e.g.,  $r=1, \dots, 8$ ) with differing opponents;

$g$  = the information condition associated with the game [ $1$  = deterministic prisoner's dilemma (DPD),  $2$  = stochastic prisoner's dilemma with full feedback (SPD-FF),  $3$  = stochastic prisoner's dilemma with partial feedback (SPD-PF)]; and

$z = 1, \dots, Z$  index the probability of loss level ( $z=1, p=0.2$ ); ( $z=2, p=0.4$ ); ( $z=3, p=0.6$ ).

The outcomes of the experiments are characterized as follows:

$Y_{itrgz} = 1$  if participant  $i$  in period  $t$  of supergame  $r$  in information condition  $g$  and in probability of loss level  $z$  chooses to invest in protection; 0 otherwise.

$L_{itrgz} = 1$  if participant  $i$  in period  $t$  of supergame  $r$  in information condition  $g$  and in probability of loss level  $z$  experiences a stochastic loss; 0 otherwise.

$M_{itrgz} = 1$  if participant  $i$  in period  $t$  of supergame  $r$  in information condition  $g$  and in probability of loss level  $z$  *could* have learned his or her counterpart's choice and 0 otherwise.

$Y_{itrgz}^c = 1$  if the counterpart  $c$  of participant  $i$  in period  $t$  of supergame  $r$  in information condition  $g$  and in probability of loss level  $z$  chooses to invest and 0 otherwise.

## 6.3 Analyzing the Prisoner's Dilemma Games using the Bayesian Model

We model the probability of investing in protection in period  $t$  as a function of a set of independent variables that includes: (i) one's loss experience in period  $t-1$ , (ii) one's own behavior in

---

<sup>9</sup> We recognize that a more general modeling framework would look at the entire path of investment decisions an individual participant made within a supergame. There is some empirical basis for focusing on just the previous round's decision. In examining experiments on coordination games, Crawford (1995) and Crawford and Broseta (1998) found that that in making a decision in period  $t$  there was a much higher weight placed on the decision in period  $t-1$  than in periods  $t-j$ ;  $j>2$ . Bostian, Holt and Smith (2007) obtained a similar result for laboratory experiments on the newsvendor problem. It is important to note that our Bayesian modeling framework is completely general and could easily incorporate decisions over a more complex set of past variables.

period  $t-1$ , (iii) whether one can learn whether one’s counterpart has invested in period  $t-1$ , (iv) the decision made by one’s counterpart  $c$  in period  $t-1$  if it can be learned, all varying by the different information conditions  $g$ , supergames  $r$ , and probability of loss conditions  $z$ . More formally we are interested in estimating the parameters of the following general model

$$\text{Probability}(Y_{itrgz} = 1) = f(L_{it-1rgz}, M_{it-1rgz}, Y_{it-1rgz}, Y_{it-1rgz}^c) \quad (1)$$

We can examine the relative importance of the variables specified in (1) using the data from our experiments and “running that data” through the lens of a Bayesian hierarchical model. The coefficients associated with each of the variables are modeled as differing from subject-to-subject (reflecting heterogeneity, e.g. some subjects are more or less influenced by their counterpart’s choices), and are assumed to be drawn from a multivariate normal distribution with general covariance matrix. By allowing for a covariance matrix among the individual-level parameters,<sup>10</sup> we can further assess, whether an individual who is influenced more by his or her counterpart’s non-investment decision is also more likely to invest following a loss. Furthermore, we also model the expected value of a subject’s coefficients as a function of both *person-level covariates* such as age, gender, race, and undergraduate major and *treatment-level covariates* such as the probability of a loss,  $z$ , and whether the subject is playing a DPD, SPD-FF or SPD-PF game,  $g$ . In this manner, we can answer the question of ‘why’ certain individuals respond in the way they do (based on individual-level characteristics) and maybe more importantly, as a function of the treatments (probability and information condition, and their interaction) that are imposed upon them.

Before laying out the model, we note two important advantages of building a Bayesian hierarchical model for our data: (1) we can properly account for the fact that players are observed multiple times and (2) we can control for confounding variables in assessing the importance of certain factors on investment decisions. As an example, a marginal analysis might show that subjects are more likely to invest after having invested in the previous round. However, suppose that investment propensity declines as the period in the game increases. Then, the effect of period is confounded with the effect of previous investment decisions. Our Bayesian hierarchical model enables us to assess the effect of previous investment decisions, holding the period fixed. A good discussion of the value of Bayesian hierarchical models is provided by Gelman, Carlin, Stern and Rubin (2004).

---

<sup>10</sup> Since the model is Bayesian, we put a prior on the covariance matrix of the multivariate normal distribution; the prior is reasonably noninformative.

In particular, we model the log odds of the probability of a participant investing (i.e.,  $\text{logit}(P_{itrgz}) = \ln(P_{itrgz}/(1 - P_{itrgz}))$ ) as a function of fixed and random effects as follows<sup>11</sup>:

Term	Explanation	Random/Fixed
$\text{logit}(P_{itrgz}) =$		
$\beta_{i1} +$	Participant-level propensity	Random
$\delta_t +$	Varying propensity by period	Fixed
$\kappa_{ig} +$	Interactions between period and information condition	Fixed
$\gamma_r +$	Varying propensity by supergame	Fixed
$\kappa_{rg} +$	Interactions between supergame and information condition	Fixed
$\beta_{i2} * Y_{it-1,rgz} +$	Effect of subject's own decision in previous period	Random
$\beta_{i3} * M_{i,t-1,rgz} * Y_{i,t-1,rgz}^c +$	Effect of counterpart investing in previous period when subject is able to learn this	Random
$\beta_{i4} * M_{i,t-1,rgz} * Y_{i,t-1,rgz} * Y_{i,t-1,rgz}^c +$	Interaction between subject's decision and counterpart's decision when subject is able to learn counterpart's decision	Random
$\beta_{i5} * L_{it-1,rgz} * (1 - Y_{i,t-1,rgz}) +$	Effect of experiencing a loss when subject did not invest in a previous round	Random
$\beta_{i6} * L_{it-1,rgz} * Y_{i,t-1,rgz} +$	Effect of experiencing a loss when subject did invest in a previous round	Random
$\beta_{i7} * L_{it-1,rgz} * (1 - Y_{i,t-1,rgz}) * Y_{i,t-1,rgz}^c * M_{i,t-1,rgz} +$	Interaction between experiencing a loss, subject's investment decision and counterpart's investment decision for when subject is able to learn counterpart's decision	Random
$\psi * \beta_{i1} * \text{Period 1 dummy}$	Additional effect of participant level propensity in period 1 to account for there being no $Y_{i(t-1)rgz}$ in period 1	Fixed

<sup>11</sup> Those effects that are individual-specific, as is standard, are modeled as random effects.

The random effects  $(\beta_{i1}, \beta_{i2}, \beta_{i3}, \beta_{i4}, \beta_{i5}, \beta_{i6}, \beta_{i7})$  are modeled as coming from a multivariate normal distribution with a mean that depends linearly on the following observed covariates:

(1) *person-level covariates*: age, gender, race, dummy variable for undergraduate, dummy variable for business major and interaction between business major and undergraduate,

(2) *treatment-level covariates*: information condition (dummy variables for deterministic condition and stochastic partial feedback condition), probability of a loss level (dummy variables for  $p=0.2$  and  $p=0.6$ ) and interactions between the information condition and the probability of a loss level (dummy variables for the combinations of deterministic condition and  $p=.2$ , deterministic condition and  $p=.6$ , stochastic partial feedback condition and  $p=.2$  and stochastic partial feedback condition and  $p=.6$ ).

In other words,

$$\begin{aligned}
 E(\beta_{ij} | i, g, z) = & \pi_{0j} + \pi_{1j} \text{age}_i + \pi_{2j} \text{gender}_i + \pi_{3j} \text{race}_i + \pi_{4j} I(i \text{ is undergraduate}) + \\
 & \pi_{5j} I(i \text{ is business major}) + \pi_{6j} I(i \text{ is undergraduate and business major}) + \\
 & \pi_{7j} I(g = \text{Deterministic (Det)}) + \pi_{8j} I(g = \text{Stochastic Partial Feedback (SPF)}) + \\
 & \pi_{9j} I(z = 0.2) + \pi_{10j} I(z = 0.6) + \pi_{11j} I(g = \text{Det}, z = 0.2) + \pi_{12j} I(g = \text{SPF}, z = 0.2) + \\
 & \pi_{13j} I(g = \text{Det}, z = 0.6) + \pi_{14j} I(g = \text{SPF}, z = 0.6) \tag{2}
 \end{aligned}$$

where  $I(x) = 1$  if condition  $x$  is true, 0 otherwise. We note, as previously mentioned, the ‘‘scientific importance’’ of equation (2) as it allows us to answer ‘whys’, i.e. what is the impact of the treatment on people’s investment propensities.

We put the following prior distribution on the parameters. For the period and supergame round effects, we used independent standard normal priors. For the covariance matrix of  $(\beta_{i1}, \beta_{i2}, \beta_{i3}, \beta_{i4}, \beta_{i5}, \beta_{i6}, \beta_{i7})$ , we used an inverse Wishart prior with 7 degrees of freedom and scale matrix  $10 * I_7$ , where  $I_7$  denotes the 7x7 identity matrix. For the coefficients on the covariates that affect the mean of  $(\beta_{i1}, \beta_{i2}, \beta_{i3}, \beta_{i4}, \beta_{i5}, \beta_{i6}, \beta_{i7})$ , we used independent standard normals.

We used the WinBUGS software (<http://www.mrc-bsu.cam.ac.uk/bugs/>) to obtain draws from the posterior distribution using Markov chain Monte Carlo (MCMC). We ran three chains of 25,000 draws each, taking the first 20,000 draws of each chain as burn-in and the last 5,000 draws of each chain as draws from the posterior distribution. We assessed convergence of the MCMC chains using Gelman-Rubin’s (1992) potential scale reduction statistic. The code and further computation details are available from the authors upon request.

## 6.4 Experimental Findings

We now use the Bayesian hierarchical model to test H5 through H9. For each of the hypotheses, we indicate the coefficient in the logit model that is used to test whether or not the experimental data provide support for it, holding all the other factors fixed. It is through the direct mapping of parameters to hypotheses that inference under the Bayesian model is made straightforward.

### 6.4.1 Testing H5

In our model, the parameter  $\beta_{i1}$  measures subject  $i$ 's propensity to invest. H5 says that there is substantial variability in  $\beta_{i1}$ . The posterior mean of the standard deviation of  $\beta_{i1}$  is 0.79 with a 95% credibility interval of (0.67, 0.92). This means that if we consider two subjects, subject 1 with  $\beta_{i1}$  one standard deviation above the mean and subject 2 with  $\beta_{i1}$  one standard deviation below the mean, then when all previous investments, losses and learning are held fixed, the odds ratio for subject 1 to invest compared to subject 2 to invest is estimated to be  $\exp(2*0.79)=4.85$ , a large effect. Thus, there is strong support for H5 – some individuals are much more likely to invest than others, holding other factors fixed.

### 6.4.2 Testing H6

The parameter  $\beta_{i2}$  measures persistence since it reflects the effect on investment in period  $t$  of having invested in period  $t-1$ , holding other factors fixed. We have thus hypothesized that the mean of  $\beta_{i2}$  is positive. We indeed find that there is strong evidence that the mean of  $\beta_{i2}$  differs between the information conditions. The posterior median (across subjects) for the mean of  $\beta_{i2}$  across the information conditions is 2.05 with a 95% credibility interval of (1.83, 2.25).<sup>12</sup> This means that for the average subject, the odds ratio for the subject to invest if he or she invested in the previous round compared to if he or she did not invest, holding all other factors fixed, is estimated to be  $\exp(2.05)=7.78$  with a 95% credibility interval of  $(\exp(1.83), \exp(2.25))=(6.23,9.49)$ . Thus, there is strong evidence for H2 that there is persistence in investment behavior in all three information conditions.

---

<sup>12</sup> The mean of  $\beta_{i2}$  across the information conditions is based on the covariate distribution in Appendix A and assumes the subjects are equally distributed across the information conditions and probability of loss conditions. For all subsequent means of random coefficients, we use the covariate distribution in Appendix A and assume the subjects are equally distributed across the relevant information conditions and probability of loss conditions.



### 6.4.3 Testing H7:

The parameter  $\beta_{i3}$  reflects how an individual's likelihood of investing in period  $t$  is impacted by learning that her counterpart invested in period  $t-1$ , holding other factors fixed. H7 says that the mean of  $\beta_{i3}$  is positive. Table 7 shows the posterior median for the mean of  $\beta_{i3}$  (across subjects) in the three information conditions.

#### [INSERT TABLE 7 HERE]

There is strong evidence that for the DPD and the SPD-FF, the average subject is more likely to invest if his or her counterpart invested in the previous round than if his or her counterpart did not. For the DPD, the median of the odds ratios among different subjects for the subject to invest if his or her counterpart invested in the previous period compared to if his or her counterpart did not is estimated to be  $\exp(1.37)=3.94$ ; for the SPD-FF, the median odds ratio is estimated to be somewhat smaller,  $\exp(0.87)=2.39$ , but still significant. For the SPD-PF, there is not strong evidence that a subject is more likely to invest if his or her counterpart invested in the previous round than if his or her counterpart did not; the median odds ratio is estimated to be only  $\exp(0.03)=1.03$ . Thus, H7 is confirmed for DPD and SPD-FF, but not for SPD-PF. The differences in the impact of learning about one's counterpart investing between SPD-FF and SPD-PF might be explained by the implicit learning in SPD-PF not being as effective as the explicit learning in DPD and SPD-FF (see hypothesis H4).

### 6.4.3 Testing the Four Versions of H8:

H8A: The odds ratio for a subject to invest in the current period if the subject invested in the previous period, the counterpart did not invest and the subject experienced a loss compared to the same conditions but the subject did not experience a loss is  $\exp(\beta_{i6})$ . Thus, H8A hypothesizes that the mean of  $\beta_{i6}$  is positive for subjects in SPD-FF. The posterior median for the mean of  $\beta_{i6}$  for subjects in SPD-FF is 0.28 with a 95% credibility interval of (-0.02, 0.58). Thus, although the point estimate supports H8A, there is only moderate evidence for H8A as the credibility interval contains zero, albeit slightly so.

H8B: The odds ratio for a subject in SPD-FF to invest in the current period if the subject did not invest in the previous period, the counterpart did invest and the subject experienced a loss compared to the same conditions but the subject did not experience a loss is  $\exp(\beta_{i5} + \beta_{i7})$ . Thus, H8B hypothesizes that the mean of  $\beta_{i5} + \beta_{i7}$  is positive for subjects in SPD-FF. The posterior median for the mean of

$\beta_{i5} + \beta_{i7}$  for subjects in SPD-FF is 0.51 with a 95% credibility interval of (0.28,0.85). Thus, there is strong evidence for H8B. The estimated median odds ratio for the effect of a loss in the situation of H8B is  $\exp(0.51)=1.67$ .

H8C: The odds ratio for a subject in SPD-FF to invest in the current period if both the subject and the counterpart did not invest and the subject experienced a loss compared to the same conditions but the subject did not experience a loss is  $\exp(\beta_{i5})$ . Thus, H8C hypothesizes that the mean of  $\beta_{i5}$  is positive for subjects in SPD-FF. The posterior median for the mean of  $\beta_{i5}$  is 0.39 with a 95% credibility interval of (0.16, 0.59). Thus, there is strong evidence for H8C. The estimated median odds ratio for the effect of a loss in the situation of H8C is  $\exp(0.39)=1.48$ .

H8D: The odds ratio for a subject in SPD-PF to invest in the current period if the subject did not invest in the previous period, the subject could not learn his or her counterpart's decision in the previous period and the subject experienced a loss compared to the same conditions but the subject did not experience a loss is  $\exp(\beta_{i5})$ . Thus, H8D hypothesizes that the mean of  $\beta_{i5}$  is positive for subjects in SPD-PF. The posterior median for the mean of  $\beta_{i5}$  for subjects in SPD-PF is 0.33 with a 95% credibility interval of (0.12, 0.54). Thus, there is strong evidence for H8D. The estimated median odds ratio for the effect of a loss in the situation of H8D is  $\exp(0.33)=1.39$ .

#### 6.4.4 Testing H9

H9 can only be tested for subjects in SPD-FF. The odds ratio for a subject to invest in the current period if the subject did not invest in the previous period, experienced a loss and the counterpart invested compared to the same conditions but the subject did not experience a loss is  $\exp(\beta_{i5} + \beta_{i7})$ . The odds ratio for a subject in SPD-FF to invest in the current period if the subject did not invest in the previous period, experienced a loss and the counterpart did not invest compared to the same conditions but the subject did not experience a loss is  $\exp(\beta_{i5})$ . H9 is hypothesizing that the former odds ratio is larger than the latter odds ratio on average. Thus, H9 is hypothesizing that the mean of  $\beta_{i7}$  is greater than zero for subjects in SPD-FF. The posterior median for the mean of  $\beta_{i7}$  for subjects in SPD-FF is 0.13 with a 95% credibility of (-0.17, 0.49). Thus, although the point estimate supports H9, there is not strong evidence for H9.

We now summarize the results of testing hypotheses H8 and H9. For the most part, we found support for the hypothesis H8 that experiencing a loss makes subjects more likely to invest in the future, holding all other conditions fixed. The only situation in which we did not find strong support for H8 was when a subject invested but his or her counterpart did not invest (H8A). In addition, we did not find strong evidence that the effect of a loss was greater when the subject's failure to invest was the sole cause of the loss as compared to when both players share some blame for the loss (H9).

#### 6.4.5 *Effects of person level covariates*

We now describe for each of the random subject coefficients  $\beta_{i1}, \dots, \beta_{i7}$  which of the six person level covariates (age, gender, race, undergraduate, business major and the interaction between undergraduate and business major), if any, had statistically significant effects on the mean of the random coefficient at a 95% confidence level.

1.  $\beta_{i1}$  (propensity to invest): None of the person level covariates had a significant effect.
2.  $\beta_{i2}$  (persistence of investment): Age had a positive effect on persistence. The mean of  $\beta_{i2}$  was estimated to increase by 0.03 for each year of age with a 95% credibility interval for this effect of (0.01, 0.06). Men were more persistent than women on average. The mean of  $\beta_{i2}$  was estimated to be 0.35 higher for men than women with a 95% credibility interval of (0.03, 0.66).
3.  $\beta_{i3}$  (increase in investment when counterpart invests): Whites increase their investment when the counterpart invests less (are less cooperative) than minorities on average. The mean of  $\beta_{i3}$  was estimated to be 0.46 lower for whites than minorities with a 95% credibility interval of (0.18, 0.75).
4.  $\beta_{i4}$  (interaction between subject's and counterpart's decision to invest): Whites have more of an interaction between subject's and counterpart's decision to invest than minorities on average. The mean of  $\beta_{i4}$  was estimated to be 0.44 higher for whites than minorities with a 95% credibility interval of (0.11, 0.79).
5.  $\beta_{i5}$  (effect of loss when subject does not invest): Undergrads respond less to losses than graduate students and non-students on average. The mean of  $\beta_{i5}$  was estimated to be 0.34 lower for undergrads with a 95% credibility interval of (0.09, 0.59).
6.  $\beta_{i6}$  (effect of loss when subject does invest): None of the person level covariates had a significant effect.
7.  $\beta_{i7}$  (additional effect of loss when subject does not invest and counterpart invests): None of the person level covariates had a significant effect.

While these covariate effects are quite suggestive, we believe that further study with a broader population is necessary.

## 7. Conclusions and Suggestions for Future Research

We now summarize the key findings from these experiments and suggest future experiments to improve our understanding of behavior in IDS situations.

### *Summary of Key Experimental Findings*

This paper provides evidence that in a two person prisoner's dilemma game individuals are much more likely to be cooperative when payoffs are deterministic (the DPD game) than when there is some chance that one will not suffer a loss, even if one does *not* invest in protection (the SPD games). A key factor behind this difference is that subjects in the SPD games respond not just to what their counterparts did but also to whether or not they suffered a loss. When a person does not invest but his or her counterpart does, the individual is less likely to reciprocate the counterpart's investment in the next period if he or she does not suffer a loss.

In comparing two stochastic prisoner's dilemma games, one with full feedback on the counterpart's behavior and the other with only partial feedback on the counterpart's behavior, we found that the overall amount of cooperation (investment) was similar. However, we did find that when subjects in the partial feedback game could learn their counterpart's actions implicitly, they reciprocated their counterparts' behavior less than in analogous situations in the full feedback game when subjects learned their counterparts' behavior explicitly.

### *Future Experiments*

The IDS-like experiments reported in this paper were based on parameters so that the only Nash equilibrium was for both players to not invest in protection if they behaved in a way that maximized expected value, even though they both would have been better off by investing. They were explicitly designed to examine prisoner-dilemma like behavior in the more realistic setting of probabilistic outcomes.

As shown in Section 3, IDS games can have two Nash equilibria so that individuals are faced with the challenges of coordination when they cannot communicate with each other. Hess, Holt and Smith (2005) have undertaken a set of laboratory experiments in such a setting with groups of four individuals, each having to decide whether to invest in protection against either an internal or external threat. They design the probabilities and payoffs so that the economically sensible decision is for all individuals to invest in protection even though there is an inferior stable NE where all players do not

invest. They show that many groups fail to coordinate their investment decisions when the game is played simultaneously but the outcomes are improved in a sequential game when participants are allowed to view the decisions of other group members as soon as they were submitted.

More generally an IDS setting opens up the possibility of tipping behavior in the spirit of Schelling (1978) and others. Future experiments could study when tipping is likely to occur if all players in the group are identical or when there is heterogeneity among the players. In either situation one could force one or more of the players to make a decision as to whether or not to invest, as Hess, Holt and Smith (2005) did in their sequential model, and determine whether other players are likely to follow suit. In theory one would expect that a player, who knows that others will observe her decision, will take an action designed to encourage others to coordinate so that everyone reaches the superior Nash equilibrium. In practice this may not be the case if the player believes that some of his counterparts will decide not to cooperate in order to maximize their own personal gains at the expense of social welfare.

Another line of experiments would be to examine individual behavior in either a simultaneous or sequential game when the probability of a loss is very low (i.e.  $p, q < .1$ ) and the loss  $L$  is very high. There is considerable empirical evidence that in such settings individuals behave as if the event will not happen to me and hence do not undertake protective measures (Kunreuther 2001). Whether one would be able to induce tipping behavior through some type of sequential game is an open question.

In the IDS experiments conducted to date, if one individual suffers a loss then the other individuals will also suffer the identical loss. Future experiments could examine behavior when there is uncertainty as to whether a counterpart will suffer a loss if you do not invest in protection (i.e.  $q < 1$ ). Another variation on this theme would be to have the size of the loss differ if it is due to your own failure to invest or from the counterpart's decision not to protect herself. For example, if a fire is caused by your failure to invest in protection then the loss to your house could be considerably higher than if the fire spread from your neighbor's house to yours.

At a prescriptive level one could also design experiments that induced one or more players to invest in protection by imposing positive economic incentives (e.g. subsidies) to encourage this actions or negative sanctions (e.g. fines) for failure to do so. Given the much larger proportion of pairs of individuals who failed to invest in protection when outcomes were uncertain than when they were deterministic it may be necessary to intervene in these ways to improve both individual and social welfare in the many IDS-like situations that we are facing in today's interdependent world.

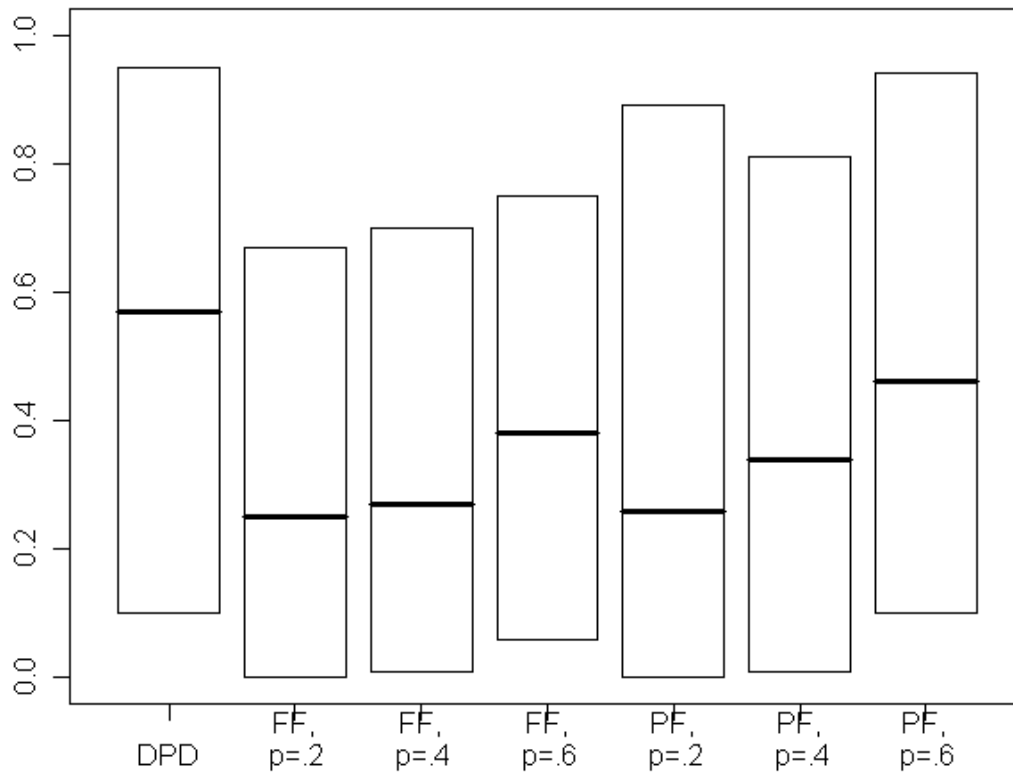
## References

- Andreoni, J. and Miller, J. (1993). Rational cooperation in the finitely repeated prisoner's dilemma: experimental evidence. *Economic Journal* 103, 570-585.
- Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books: New York.
- Axelrod, R. and Dion, D. (1988). The further evolution of cooperation. *Science* 242, 1385-1389.
- Axelrod, R. and Hamilton, W.D. (1981). The evolution of cooperation. *Science* 211, 1390-1396.
- Bendor, J. (1987). In good times and bad: reciprocity in an uncertain world. *American Journal of Political Science* 31, 531-538.
- Bendor, J. (1993). Uncertainty and the evolution of cooperation. *American Journal of Political Science* 31, 531-538.
- Bendor, J., Kramer, R.M. and Stout, S. (1991). When in doubt... cooperation in a noisy prisoner's dilemma. *Journal of Conflict Resolution* 35, 691-719.
- Bereby-Meyer, Y. and Roth, A.E. (2006). The speed of learning in noisy games: partial reinforcement and the sustainability of cooperation. *American Economic Review* 96, 1029-1042.
- Bostian, A., Holt, C. and Smith, A. (2007) The Newsvendor "Pull-to-Center Effect": Adaptive Learning in a Laboratory Experiment, Working Paper, University of Virginia (May).
- Cooper, R., Dejong, D.V., Forsythe, R. and Ross, T.W. (1996). Cooperation without reputation: experimental evidence from the prisoner's dilemma games. *Games and Economic Behavior* 12, 187-218.
- Crawford, V. (1995) Adaptive Dynamics in Coordination Games *Econometrica* 63, 103-143.
- Crawford, V. and Broseta, B. (1998) What Price Coordination? The Efficiency-Enhancing Effect of Auctioning the Right to Buy *American Economic Review* 88, 198-225.
- Dawes, R.M. (1980). Social dilemmas. *Annual review of psychology* 31, 169-193.
- Donninger, C. (1986). Is it always efficient to be nice? A computer simulation of Axelrod's computer tournament. In *Paradoxical Effects of Social Behavior*, pp. 123-134, A. Diekmann and P. Mitter eds. Physica-Verlag: Heidelberg.
- Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2004). *Bayesian Data Analysis, Second Edition*. Chapman and Hall: Boca Raton.
- Gelman, A. and Rubin, D.B. (1992), Inference from iterative simulation using multiple sequences (with discussion), *Statistical Science* 7, 457-511.

- Hardin, G. (1968). The tragedy of the commons. *Science* 162 : 1243-1248.
- Hauk, E. and Nagel, R. (2001). Choice of partners in multiple two-person prisoner's dilemma games: an experimental study. *Journal of Conflict Resolution* 45, 770-793.
- Heal, G. and Kunreuther, H. (2005). You only die once: interdependent security in an uncertain world in H.W. Richardson, P. Gordon and J.E. Moore II, eds., *The Economic Impacts of Terrorist Attacks* Cheltenham, UK: Edward Elgar.
- Hess, R. Holt, C. and Smith, A. (2005) "Coordination of Strategic Responses to Security Threats: Laboratory Evidence," Paper prepared for the Coordination Workshop in Honor of Ray Battalio, Texas A&M.
- Kreps, D., Milgrom, P., Roberts, J. and Wilson, R. (1982). Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory* 27, 245-252.
- Kunreuther, H. and Heal, G. (2003). Interdependent security. *Journal of Risk and Uncertainty* 26, 231-249.
- Kunreuther, H. (2001) "Protective Decisions: Fear or Prudence?" in S. Hoch and H. Kunreuther (ed.) *Wharton on Making Decisions* New York: Wiley.
- Kunreuther, H. (2006) "Comprehensive Disaster Insurance: Has its Time Come?" in *Risk and Disaster: Lessons from Hurricane Katrina*. University of Pennsylvania Press, Philadelphia, PA.
- Liang, K.-Y. and Zeger, S.L. (1986), Longitudinal data analysis using generalized linear models. *Biometrika* 73, 13-22.
- Molander, P. (1985). The optimal level of generosity in a selfish, uncertain environment. *Journal of Conflict Resolution* 29, 611-618.
- Mueller, U. (1987). Optimal relation for optimal cooperation. *Journal of Conflict Resolution* 31, 692-724.
- Ostrom, E., Dietz, T., Dolšak, N., Stern, P.C., Stonich, S. and Weber, E.U., eds. *The Drama of the Commons* Washington, D.C.: The National Academy Press.
- Schelling, T.J. (1978). *Micromotives and macrobehavior*. W.W. Norton & Company: New York.
- Selten, R. and 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.

**Figure 1**

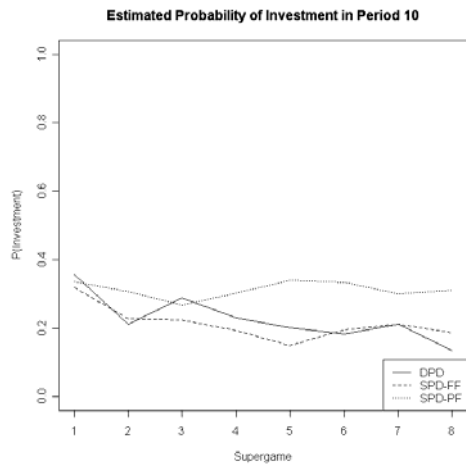
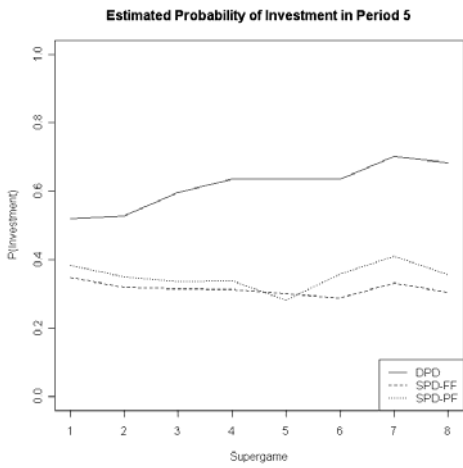
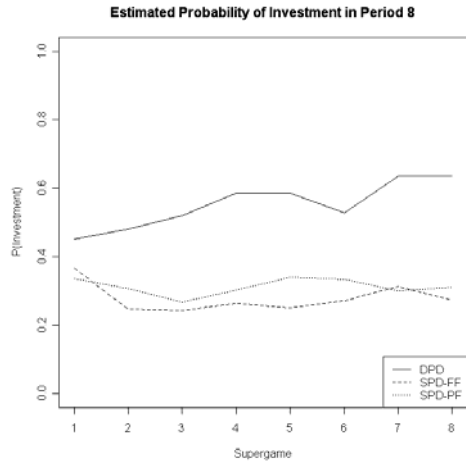
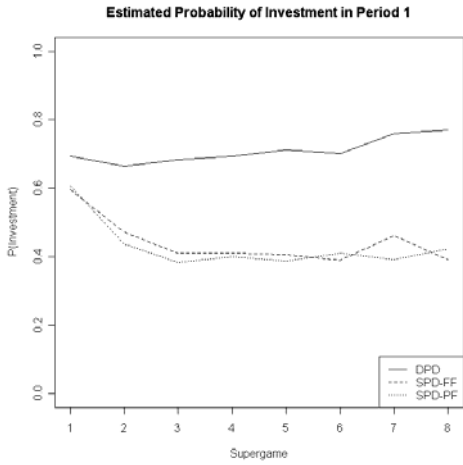
**Mean, .05 and .95 Quantiles of Investment in Different Conditions**



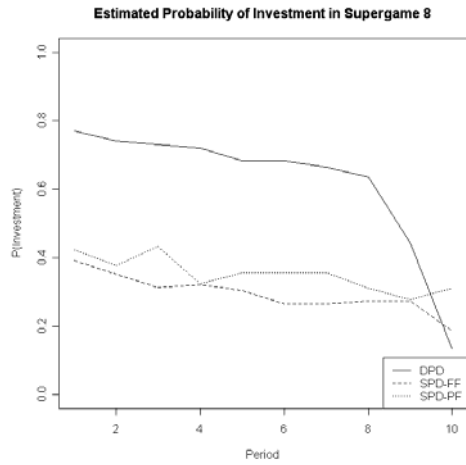
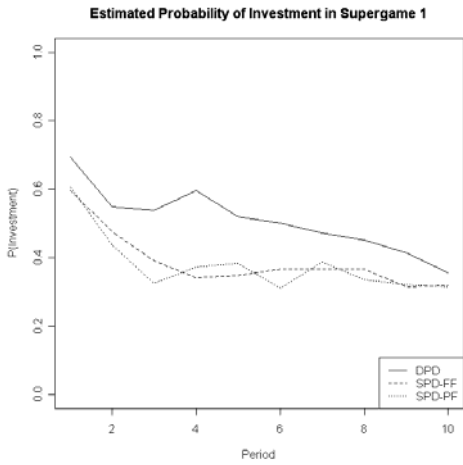
The ends of the boxes show the .05 and .95 quantiles of the distribution of subject investment proportions in a condition, where the investment proportion for a given subject is computed using all supergames the subject played. The dark line in the middle of the box shows the mean investment across subjects in the condition.

**Figure 2**





**Figure 3**



**Table 1. Expected Returns Associated with Investing and Not Investing in Protection**

	<i>Individual 2 ( A<sub>2</sub> )</i>		
	<b>I</b>		<b>NI</b>
<i>Individual 1 ( A<sub>1</sub> )</i>	<b>I</b>	$Y-c, Y-c$	$Y-c-qL, Y-pL$
	<b>NI</b>	$Y-pL, Y-c-qL$	$Y-[pL + (1-p) qL], Y-[pL+ (1-p)qL]$

**Table 2. Scenarios related to Decisions in Period t and whether or not learning is possible in the Stochastic Partial-Feedback Condition:**

Scenario	Player 1 Decision	Player 2 Decision	Color	Loss or No Loss?	Player 1 can learn decision of Player 2?
1	I	I	Red	No Loss	Yes
2	I	I	Orange	No Loss	No
3	I	I	Green	No Loss	No
4	I	NI	Red	Loss	Yes
5	I	NI	Orange	No Loss	No
6	I	NI	Green	No Loss	No
7	NI	I	Red	Loss	No
8	NI	I	Orange	No Loss	Yes
9	NI	I	Green	No Loss	No
10	NI	NI	Red	Loss	No
11	NI	NI	Orange	Loss	Yes
12	NI	NI	Green	No Loss	No

**Table 3. Number of Individuals in Each Experimental Condition**

Condition	Number of Supergames	Number of Individuals
DPD	84	104
SPD-FF (all)	140	210
p=0.2	40	54
p=0.4	40	76
p=0.6	60	80
SPD-PF (all)	140	206
p=0.2	40	52
p=0.4	50	96
p=0.6	50	58
<b>TOTAL</b>	<b>324</b>	<b>520</b>

**Table 4. Percentage of Individuals Investing in Protection in the Three Conditions**

Condition (Loss)	Total Decisions	Total I Decisions	I/Total
DPD (L=10)	8800	5039	0.57
SPD-FF (all)	14800	4626	0.31
P=0.2 (L=50)	4320	1094	0.25
P=0.4 (L=25)	4440	1211	0.27
P=0.6 (L=19)	6040	2321	0.38
SPD-PF (all)	13600	4753	0.35
P=0.2 (L=50)	4140	1092	0.26
P=0.4 (L=25)	5400	1809	0.34
P=0.6 (L=19)	4060	1852	0.46
All P=0.2 (L=50)	8460	2186	0.26
All P=0.4 (L=25)	9840	3020	0.31
All P=0.6 (L=19)	10100	4173	0.41

**Table 5 Confidence Intervals for Contrasts in Proportion of Investment Between Conditions**

Contrast	Estimated Difference	95% CI	
(SPD-FF, p=.2)-DPD	-0.34	(-0.41,-0.27)	R1
(SPD-PF, p=.2)-DPD	-0.30	(-0.37,-0.23)	R2
(SPD-FF, p=.4)-DPD	-0.29	(-0.40,-0.18)	R3
(SPD-PF, p=.4)-DPD	-0.25	(-0.36,-0.14)	R4
(SPD-FF, p=.6)-DPD	-0.18	(-0.29,-0.07)	R5
(SPD-PF, p=.6)-DPD	-0.13	(-0.02,-0.25)	R6
(SPD-FF, p=p*)- (SPD-PF, p=p*) for p*=(.2,.4.6)	-0.04	(-0.09,0.00)	R7
p=.6 – p=.2 for fixed SPD game	0.16	(0.10,0.22)	R8
p=.6 – p=.4 for fixed SPD game	0.12	(0.06,0.17)	R9
p=.4 – p=.2 for fixed SPD game	0.05	(-0.01,0.10)	R10

**Table 6**

Previous Decision by Subject	Previous Decision by Counterpart	Investment Proportion in SPD-FF	Investment Proportion in SPD-PF when counterpart's previous decision is learned
I	I	0.88	0.79 R1
NI	I	0.21	0.15 R2
I	NI	0.44	0.40 R3
NI	NI	0.08	0.19 R4

**Table 7**

Information Condition	Posterior Median of $\beta_{i3}$	95% Credibility Interval for $\beta_{i3}$
DPD	1.37	(0.99, 1.82)
SPD-FF	0.87	(0.58, 1.12)
SPD-PF	0.03	(-0.45, 0.43)

**Appendix A: Demographic Information about the Subject Pool**

Total number of subjects: 520

(Note that some subjects did not identify their gender, age, and/or ethnicity.)

Male: 226

Female: 293

Asian or Pacific Islander: 205

Black, not of Hispanic Origin: 37

Hispanic: 24

White, not of Hispanic Origin: 229

Other: 24

17 yrs old: 6

18: 84

19: 101

20: 92

21: 71

22: 38

23: 20

24: 9

25: 7

older than 25: 88

Undergraduate Student, Business: 164

Undergraduate Student, Arts & Sciences: 158

Undergraduate Student, Engineering: 65

Undergraduate Student, Nursing: 6

Graduate Student, Arts & Sciences: 17

Graduate Student, Engineering: 10

Graduate Student, Other (Med, Law, etc.): 38

Other/Non-Student: 63

## Appendix B. Instructions to Subjects

### 1. Instructions Presented to Subjects in the Deterministic Prisoner's' Dilemma Condition

*[Opening Instruction Page:]*

This is a game in which the outcomes of your decisions depend not only on what *you* do, but also on what your *counterpart* does.

You will be paired with another person in the room whose identity is not known to you. In each of 10 rounds, you and your counterpart will independently make a decision about whether or not to invest funds to avoid a financial loss from a negative event.

- If both of you choose to INVEST then it will cost each of you 12 talers, but neither of you will experience a financial loss from a negative event.
- If one of you INVESTS and the other does NOT INVEST, it will cost the one who INVESTS 12 Talers. In addition, both people will suffer an equal financial loss of 10 Talers from a negative event.
- If both of you choose to NOT INVEST, then each of you suffers an equal financial loss of 16 Talers from a negative event.

Below is the summary of the possible outcomes:

		Your Counterpart	
		INVEST	NOT INVEST
You	INVEST	<ul style="list-style-type: none"> <li>• You lose <b>12</b>.</li> <li>• Your counterpart loses <b>12</b>.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose <b>22</b>.</li> <li>• Your counterpart loses <b>10</b>.</li> </ul>
	NOT INVEST	<ul style="list-style-type: none"> <li>• You lose <b>10</b>.</li> <li>• Your counterpart loses <b>22</b>.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose <b>16</b>.</li> <li>• Your counterpart loses <b>16</b>.</li> </ul>

*[Figure B1]*

You and your counterpart are each given 300 Talers (10 talers = \$1) before you start making decisions. You will not know the decision your counterpart has made until the end of the round. Before the start of the next round you will be given feedback on what each of you did and the status of your assets.

One pair will be chosen at random to receive the dollar equivalent of the talers they have at the end of the game (10 talers = \$1).

[Each subject sees the following screen before making his/her decision for the first round:]

**Round 1**

Your counterpart is also deciding.

Please choose from the following two options:

I will Invest.

I will NOT Invest.

Payoff Table for All Rounds:

		Your Counterpart	
		INVEST	NOT INVEST
You	INVEST	<ul style="list-style-type: none"> <li>• You lose 12.</li> <li>• Your counterpart loses 12.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose 22.</li> <li>• Your counterpart loses 10.</li> </ul>
	NOT INVEST	<ul style="list-style-type: none"> <li>• You lose 10.</li> <li>• Your counterpart loses 22.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose 16.</li> <li>• Your counterpart loses 16.</li> </ul>

[Figure B2]

[After each person in the pair has made a decision, subjects see their decisions and payoffs highlighted in a table:]

**Round 1**

You have chosen to Invest, but your counterpart has chosen to Not Invest, as indicated in blue in the table below.

		Your Counterpart	
		Invest	Not Invest
You	Invest	<ul style="list-style-type: none"> <li>• You lose 12 talers.</li> <li>• Your counterpart loses 12 talers.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose 22 talers.</li> <li>• Your counterpart loses 10 talers.</li> </ul>
	Not Invest	<ul style="list-style-type: none"> <li>• You lose 10 talers.</li> <li>• Your counterpart loses 22 talers.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose 16 talers.</li> <li>• Your counterpart loses 16 talers.</li> </ul>

You Lose 22 Talers and your counterpart loses 10 Talers.

[Figure B3]

[In subsequent rounds each subject sees a history report of past decisions in the supergame. The subject sees this on the decision screen:]

### Round 5

Amounts lost in previous rounds:

Round	Your Decision	Your Outcome	Your Ending Balance	Counterpart's Decision	Counterpart's Outcome	Counterpart's Ending Balance
1	Invest	-22	278	Not Invest	-10	290
2	Not Invest	-16	262	Not Invest	-16	274
3	Not Invest	-16	246	Not Invest	-16	258
4	Invest	-22	224	Not Invest	-10	248

Your counterpart is also deciding.

Please choose from the following two options:

- I will Invest.
- I will NOT Invest.

Submit

Payoff Table for All Rounds:

		Your Counterpart	
		INVEST	NOT INVEST
You	INVEST	<ul style="list-style-type: none"> <li>• You lose 12.</li> <li>• Your counterpart loses 12.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose 22.</li> <li>• Your counterpart loses 10.</li> </ul>
	NOT INVEST	<ul style="list-style-type: none"> <li>• You lose 10.</li> <li>• Your counterpart loses 22.</li> </ul>	<ul style="list-style-type: none"> <li>• You lose 16.</li> <li>• Your counterpart loses 16.</li> </ul>

[Figure B4]

**2a. Instructions Presented to Subjects in the Stochastic Full-Feedback Condition (for p=0.2):**

*[Opening Instruction Page:]*

This is a game in which the outcomes of your decisions depend not only on what *you* do, but also on what your *counterpart* does.

You will be paired with another person in the room whose identity is not known to you. In each of 10 rounds, you and your counterpart will independently make a decision about whether or not to invest funds to avoid a financial loss from a random negative event.

Financial losses will be measured in a fictitious currency called "Talers".

- If both you and your counterpart choose to INVEST, then the investment cost to each of you is 12 talers.
- If you INVEST and your counterpart does NOT INVEST, then there is a 20% chance that your counterpart will lose 50 talers and you will lose 62 talers; and there is an 80% chance that your counterpart will lose 0 talers and you will lose 12 talers.
- If you do NOT INVEST and your counterpart INVESTS, then there is a 20% chance that your counterpart will lose 62 talers and you will lose 50 talers; and there is an 80% chance that your counterpart will lose 12 talers and you will lose 0 talers.
- If both you and your counterpart choose to NOT INVEST, then each of you has a 36% chance of losing 50 talers, and a 64% chance of losing 0 talers.

Probabilistic outcomes will be determined by the following Random Number Generator, where it is equally likely that any number between 1 and 100 is chosen.

For Example: If the Random Number generated is 6, then 6 will flash as follows:

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

**Red: Lose if at least one person has not invested**

**Orange: Lose if both people have not invested**

**Green: No loss**

Below is a summary of the possible outcomes:



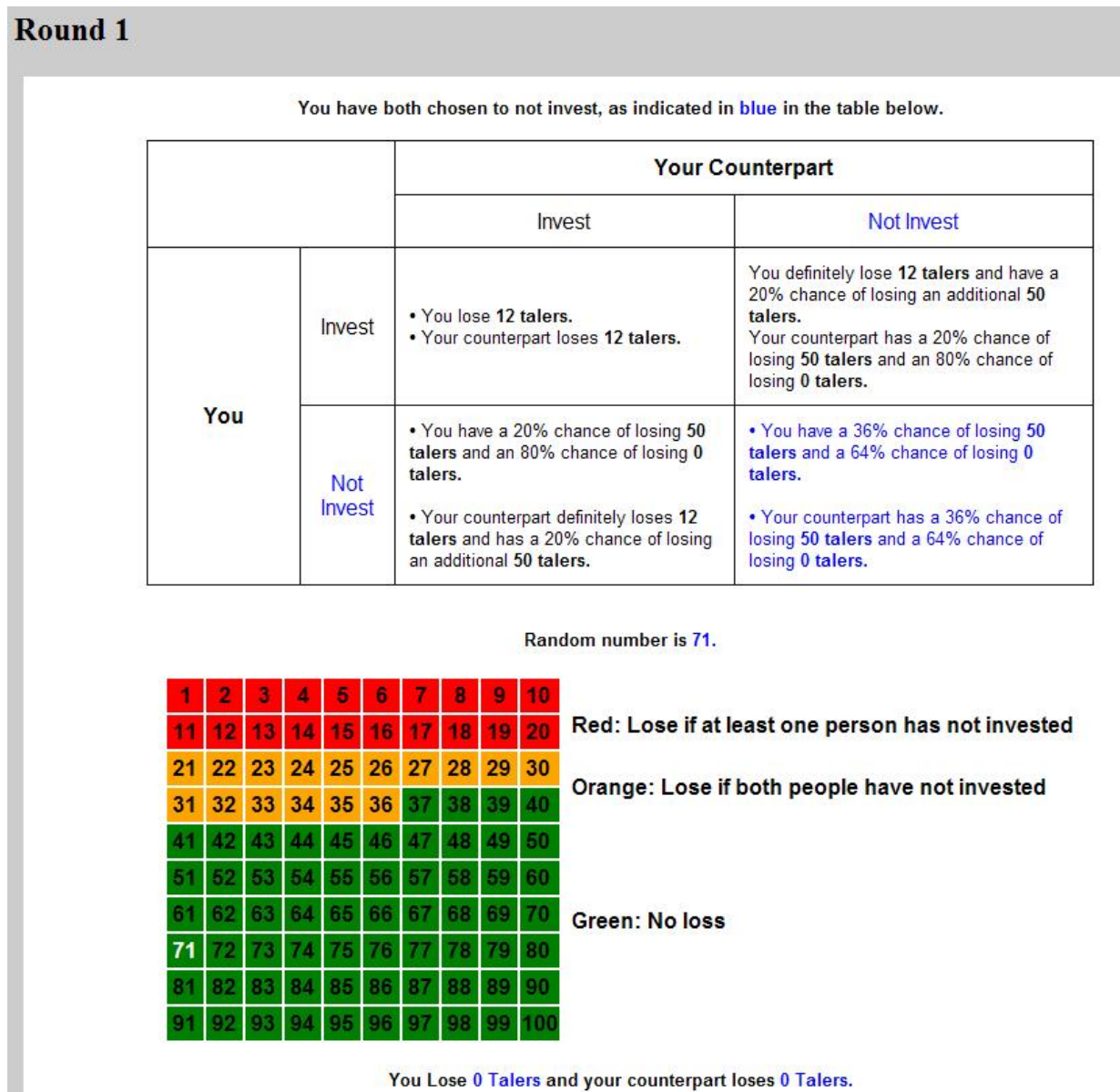
		Your Counterpart	
		INVEST	NOT INVEST
You	INVEST	<ul style="list-style-type: none"> <li>You lose <b>12 talers</b>.</li> <li>Your counterpart loses <b>12 talers</b>.</li> </ul>	<ul style="list-style-type: none"> <li>You definitely lose <b>12 talers</b> and have a 20% chance of losing an additional <b>50 talers</b>.</li> <li>Your counterpart has a 20% chance of losing <b>50 talers</b> and an 80% chance of losing <b>0 talers</b>.</li> </ul>
	NOT INVEST	<ul style="list-style-type: none"> <li>You have a 20% chance of losing <b>50 talers</b> and an 80% chance of losing <b>0 talers</b>.</li> <li>Your counterpart definitely loses <b>12 talers</b> and has a 20% chance of losing an additional <b>50 talers</b>.</li> </ul>	<ul style="list-style-type: none"> <li>You have a 36% chance of losing <b>50 talers</b> and a 64% chance of losing <b>0 talers</b>.</li> <li>Your counterpart has a 36% chance of losing <b>50 talers</b> and a 64% chance of losing <b>0 talers</b>.</li> </ul>

*[Figure B5]*

You and your counterpart are each given 300 Talers before you start making decisions. You will not know the decision your counterpart has made until the end of each round. Before the start of the next round you will be given feedback on what each of you did, whether or not a negative event occurred and the status of your assets.

One pair will be chosen at random to receive the dollar equivalent of the talers they have at the end of the game (10 talers = \$1).

[The payoff screen (what each pair sees after making their decisions for a round) looks like this:]



[Figure B6]

[The “Results of previous rounds” history table on the decision screen looks like this:]

Round	Negative Event	Your Decision	Your Outcome	Your Ending Balance	Counterpart's Decision	Counterpart's Outcome	Counterpart's Ending Balance
1	No	Not Invest	-0	300	Not Invest	-0	300
2	Yes	Invest	-62	238	Not Invest	-50	250
3	No	Not Invest	-0	238	Invest	-12	238
4	No	Not Invest	-0	238	Not Invest	-0	238

[Figure B7]

**2b. Instructions Presented to Subjects in the Stochastic Full-Feedback Condition (for  $p=0.4$ ):**

*[Opening Instruction Page is analogous to the  $p=0.2$  condition, but with the following changes to the color-grid and the payoff matrix:]*



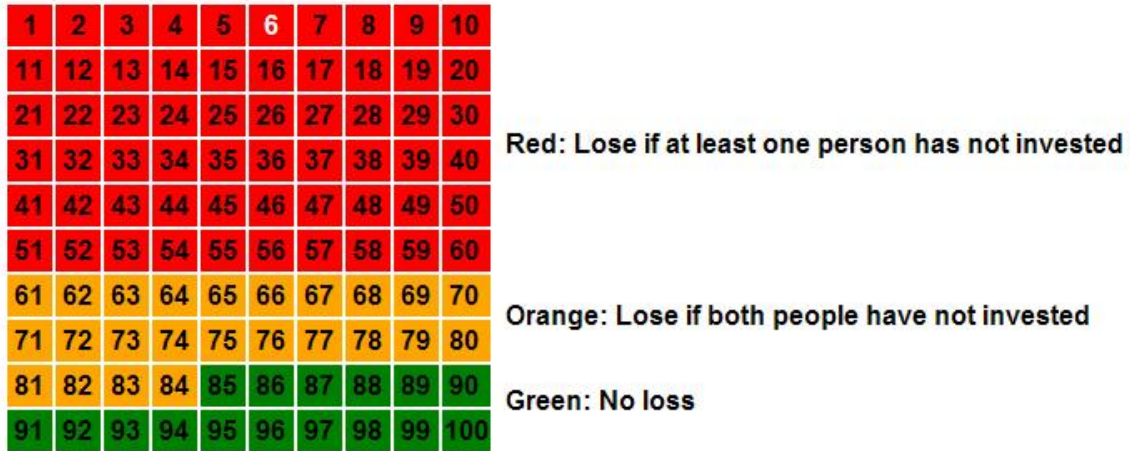
		Your Counterpart	
		INVEST	NOT INVEST
You	INVEST	<ul style="list-style-type: none"> <li>You lose <b>12 talers</b>.</li> <li>Your counterpart loses <b>12 talers</b>.</li> </ul>	<ul style="list-style-type: none"> <li>You definitely lose <b>12 talers</b> and have a 40% chance of losing an additional <b>25 talers</b>.</li> <li>Your counterpart has a 40% chance of losing <b>25 talers</b> and a 60% chance of losing <b>0 talers</b>.</li> </ul>
	NOT INVEST	<ul style="list-style-type: none"> <li>You have a 40% chance of losing <b>25 talers</b> and a 60% chance of losing <b>0 talers</b>.</li> <li>Your counterpart definitely loses <b>12 talers</b> and has a 40% chance of losing an additional <b>25 talers</b>.</li> </ul>	<ul style="list-style-type: none"> <li>You have a 64% chance of losing <b>25 talers</b> and a 36% chance of losing <b>0 talers</b>.</li> <li>Your counterpart has a 64% chance of losing <b>25 talers</b> and a 36% chance of losing <b>0 talers</b>.</li> </ul>

*[Figure B8]*

*[Other screens are formatted analogously to those shown above for  $p=0.2$ .]*

**2c. Instructions Presented to Subjects in the Stochastic Full-Feedback Condition (for p=0.6):**

*[Opening Instruction Page is analogous to the p=0.2 and p=0.4 conditions, but with the following changes to the color-grid and the payoff matrix:]*



		Your Counterpart	
		INVEST	NOT INVEST
You	INVEST	<ul style="list-style-type: none"> <li>You lose <b>12 talers</b>.</li> <li>Your counterpart loses <b>12 talers</b>.</li> </ul>	<ul style="list-style-type: none"> <li>You definitely lose <b>12 talers</b> and have a 60% chance of losing an additional <b>19 talers</b>.</li> <li>Your counterpart has a 60% chance of losing <b>19 talers</b> and a 40% chance of losing <b>0 talers</b>.</li> </ul>
	NOT INVEST	<ul style="list-style-type: none"> <li>You have a 60% chance of losing <b>12 talers</b> and a 40% chance of losing <b>0 talers</b>.</li> <li>Your counterpart definitely loses <b>12 talers</b> and has a 60% chance of losing an additional <b>19 talers</b>.</li> </ul>	<ul style="list-style-type: none"> <li>You have an 84% chance of losing <b>25 talers</b> and a 16% chance of losing <b>0 talers</b>.</li> <li>Your counterpart has an 84% chance of losing <b>19 talers</b> and a 16% chance of losing <b>0 talers</b>.</li> </ul>

*[Figure B9]*

*[Other screens are formatted analogously to those shown above for p=0.2.]*

### **3. Instructions Presented to Subjects in the Stochastic Partial-Feedback Condition:**

*[The Stochastic Partial-Feedback condition differs from the Full-Feedback as follows:]*

*[1. Opening Instructions in the Partial-Feedback condition are the same as in the Full-Feedback condition except that the penultimate paragraph reads as follows:]*

You and your counterpart are each given 300 Talers before you start making decisions. You will not know the decision your counterpart has made. Before the start of the next round you will be given feedback on what you did and the status of your assets.

*[2. The payoff screen (B6) does not indicate counterpart's decision or talers lost.]*

*[3. The history table ("Results of previous rounds") does not show columns relating to counterpart information.]*