A simple principal-agent experiment for the classroom.


Abstract:
In this note we describe a simple, flexible and instructive moral hazard experiment. It can be used in a variety of classes, including principles classes, to illustrate the basic incentive conflicts in principal-agent interactions, the importance of information, and the power of reputational enforcement.

Full Text: COPYRIGHT 1997 Western Economic Association International

Interest in the use of classroom experiments has increased dramatically in the past decade. (1) As we were surveying available demonstrations for a booklet, Ortmann and Colander [1995], we found that there were a number of classroom experiments readily available to demonstrate pricing institutions such as auctions, and symmetric game problems of the prisoner's dilemma variety which are relevant to public goods/externalities/oligopoly/cartel situations. However, there is a relative dearth of classroom experiments illustrating asymmetric game problems.

The most well-known of these asymmetric game problems are moral hazard, principal-agent games - games in which one of the players (the agent) is informed about a key aspect of the game while the other (the principal) is not. We found the lack of classroom experiments involving such games surprising and unsatisfying because they have become a prominent staple in many textbooks as in Stiglitz [1993]; Colander [1995]; Carlton and Perloff [1995]; and Mishkin [1995].

This note addresses the imbalance. We describe a simple, flexible and instructive moral hazard experiment for use in a variety of classes, including introductory courses. The experiment can be used to illustrate the importance of information, and the power of reputational enforcement in principal-agent interactions. Such issues are the key for an understanding of modern theories of the firm, such as Holmstrom and Tirole [1989], Kreps [1990a; 1990b], Stiglitz [1993], Carlton and Perloff [1995], and Mishkin [1995], and the role of market forces in assuring contractual performance as in Klein and Leffler [1981]. We begin by discussing prisoner's dilemma and principal-agent games and their relevance to economics. Then we describe the experiment and its likely results.

1. PRISONER'S DILEMMA AND PRINCIPAL-AGENT GAMES

Prisoner's dilemma games are symmetric game problems. They are of interest because they capture situations in which the collectively optimal decision will not (necessarily) be achieved through individual optimization as in Axelrod [1984]. Players' sets of conflicting choices are interchangeable; individuals
face identical dilemmas. Specifically, both prisoners are presented with the option to confess, or not to confess. Prisoners face the dilemma that their self-interest suggests that they ought to confess; yet, if they both follow their self-interested individual optimization they will be worse off than if they both cooperated. Cartel and free-rider problems fall within this type of symmetric game problem.

Moral hazard, principal-agent interactions are asymmetric game problems. They can be conceptualized as asymmetric prisoner's dilemma games as in Rasmusen [1989]. The asymmetry results from the different set of choices that the principal and the agent face. For example, consider a buyer (a principal) who takes his stereo for repairs. The repair person (the agent) diagnoses the source of the problem and promises to get a high quality part to fix it. The buyer has to decide if he wants to trust the repair person, both as regards her diagnosis and the promise to use high quality parts. If he decides not to trust her, he can get both a second opinion beforehand and have the parts checked afterwards. Both activities would require additional time and cost and may therefore not be desirable. The dilemma is obvious: the agent faces the temptation to renege on a promise that would make agent and principal together better off. (This is the moral hazard aspect of the principal-agent game presently discussed.) The principal, knowing this, is confronted with the dilemma of trusting or not trusting the agent. What makes this asymmetric "dilemma of trust" interesting and relevant to many decisions is that, similar to the symmetric prisoner's dilemma game, the agents' individual and collective rankings of outcomes differ.

In the simplest version of this asymmetric game the two players can choose between two actions each so that there are four possible outcomes. To formally specify the dilemma of trust, the game must be parameterized by assigning values to the outcomes. For the parameterization to reflect the moral hazard problem, the values chosen must be sufficient to give the agent an incentive to choose the individually advantageous, rather than the collectively desirable choice.

Following Rasmusen [1989] and Kreps [1990a; 1990b], Figure 1 shows one possible parameterization for the repair person-buyer, high quality-low quality problem described above. In this matrix, the buyer's (principal's) options are listed horizontally. He can either inspect or not inspect. The repair person's (agent's) options are listed vertically - she can either provide a high or low quality repair. The valuations of the outcomes are listed in the cells - the repair person's are listed first, and the buyer's are listed second. The values are chosen so that the joint payoffs in the upper left corner are greater than those in the lower left corner, which in turn are higher than those in the lower right corner; this parameterization captures the moral hazard aspects of the principal-agent problem because by moving from the upper left corner to the lower left the agent can make herself better off, while the principal is made worse off.

Thus, if the buyer does not inspect and the repair person provides a high quality repair, they both receive a benefit of 1. This can be thought of as the cooperative outcome. If the repair person provides a low quality repair, she doubles her benefit to 2 while the buyer is worse off; in fact, the buyer incurs a loss of 1 because he pays the price of high quality service without getting it. The buyer, knowing the repair person's temptation, thus has a strong incentive to inspect. However, the payoff of both participants is 0 if he should do so.

The parameterization is based on the following assumptions. The value of a high quality repair to the buyer is 3. The value of a low quality repair is 1. The price of the repair is 2. Thus, if he does not inspect, the buyer's net benefit of a high quality repair is 1, and of a low quality repair is -1. These are the second numbers in the first column.

The repair person's cost of a high quality repair is 1 and the cost of a low quality repair is 0, giving her a net benefit of 1 for a high quality repair and for 2 for a low quality repair if the buyer does not inspect. These are the first numbers in the first column.
The costs of inspection are assumed to be 1 for the seller, and 1 for the buyer. (Our intuitive rationale for this division is that the seller must stay there and watch as the buyer disassembles the unit, so it has a cost in time for both of them.) This means that if the buyer chooses to inspect and the seller provided a high quality repair, the payoff in the first row, second column reduces to 0 for both of them. We further assume that if the buyer chooses to inspect and the seller provided a low quality repair, the seller will have to provide a high quality repair to the buyer at a cost of 1, as promised. This accounts for the repair person's payoff of 0 in the second row, second column. Her revenues of 2 are offset by the costs of repair and the inspection. Similarly, the benefit of 3 of a high quality repair and its costs (price = 2 and inspection = 1) add up to net benefits of 0 for the buyer, accounting for the buyer's payoff of 0 in the second column, second row.

We chose this parameterization for its pedagogical simplicity. It makes it easier for students to see the moral hazard dimension of the problem. Other parameterizations are possible, and professors in intermediate and upper classes may want to encourage students to explore the rationale and robustness of the parameterization.(2,3)

The asymmetric game problem presented here has no obvious solution. To determine a solution, we must make additional assumptions. Two standard assumptions used in game theory are individual (self-interested) rationality and common knowledge of the payoff matrix.

For a one-shot game, these assumptions lead to the following chain of reasoning. If the buyer and the repair person were to agree on playing the upper left corner, the repair person could make herself better off if she switched strategies and ended in the lower left corner. If the buyer knows the payoff matrix and anticipates the reasoning of the agent, then he can make himself better off by switching strategies so that he ends up in the lower right corner. This outcome is the standard game-theoretic prediction for the one-shot version of this game.

It is not the only possible outcome of the game. In a repeated game, the intuition behind the game will be different. What changes is the intuition involving the cost and benefit of offering trust. In a repeated game the players can invest in trust, and choose the cooperative strategy at a potential cost now, but with a potential gain in the future. More precisely, in every period an agent has to trade off the gains from choosing low quality in that period against the gains from choosing high quality in both this and all future periods. This intertemporal trade-off crucially depends on the agent's discount rate for future payoffs. The more potential trust there is, the more likely the individuals are to choose a cooperative solution.(4)

Most introductory courses are not ready for the formal presentations of games, even simple ones. However, students can easily understand the general ideas, and, once a professor has presented these ideas, he or she is in a position to discuss the intuition behind many modern macro policy debates involving credibility as well as new approaches to economic problems (for instance, efficiency wages).

Even in those introductory courses where the professor believes that the class can benefit from formal theory, we recommend that only a heuristic discussion, with examples, be presented before the experimental demonstration is conducted. In intermediate classes, the formal structure can be presented so that the classroom experiment itself provides the impetus for the students to learn about it. Regardless of the level of the class, we suggest starting the game with a discussion of a very specific principal-agent problem either in the form of a story such as the one we told or in the form of a matrix such as the one presented in Figure 1.

II. THE PRINCIPAL-AGENT EXPERIMENTAL DEMONSTRATION

This principal-agent experiment is a simple representation of the asymmetric dilemma of trust problem discussed above.

The physical requirements of the experiments are the following:

1. An even number of participants. The only restriction on the number of participants is the instructor's budget constraint. (In large classes, it may make sense to select a group of between 10 to 20 students and have the others observe.)

2. Two quarters per participant. This is the maximum amount of money needed. Chances are that the actual cost of this experimental demonstration will be lower. (If you plan to play several rounds, multiply the maximum amount needed by the number of rounds you intend to play.)

3. As many pieces of paper as participants; half of them in one color (for instance red), the other half in another (for instance blue). (If you plan on several rounds of play, multiply the number of pieces of paper accordingly.)

Begin the demonstration by presenting the case described above: the agent is a seller of a good (hi-fi stereo or computer repair), supplying either high or low quality repair. The principal is the buyer who must either verify the quality (at a cost) or trust the seller's promise to deliver high quality. The next step is to discuss the dilemma the players face and how the collectively optimal decision will not (necessarily) be achieved by individual optimization.

Having described the situation to the students, select the participants and ask them to come to the front of the room. Give one half of the participants red papers and the other half blue ones. Tell the students that those with red papers are buyers and those with blue papers are sellers. Inform the students with blue papers that they are supplying repairs that can be of high or low quality. The students with red papers are buyers; they can either trust the seller's diagnosis and promise to use a high quality part, or they can have it checked out.

Each participant (both buyers and sellers) should be given an initial one-quarter endowment. Tell sellers that they should write on a piece of paper whether they want to sell a high or low quality part. Tell buyers that they should write on a piece of paper whether they will be checking the quality of the part.

Tell them that they will receive the following payments: (1) if the seller ends up providing high quality and the buyer ends up not inspecting then they will both receive another quarter; (2) if the seller ends up providing low quality and the buyer ends up not inspecting, the seller will receive an additional quarter and will also receive the buyer's quarter; and (3) if the buyer has chosen to inspect, they both will get to keep their quarter regardless of the seller's decision.

Instruct participants to walk randomly around the room until each buyer is paired with a seller, and vice versa. Once paired, they may talk with each other about the optimal strategy combination, but they must write down their choice privately. Instruct participants to fold the papers after they have made their decisions, so that the others cannot see what they have chosen. Instruct students to write down their names (initials) on the pieces of paper. Finally, ask them to hand in the folded pieces of paper in pairs - one paper of each color.

After participants have returned to their seats, record the outcomes one at a time. Pay out or re-distribute the quarters as promised, and then discuss the results with the class. For example, ask the students why they made the choices they did.

The experiment can be repeated, and a choice made as to whether the track records of participants are to be revealed or kept secret. (See Tullock [1985] and Frank [1988] for inspiration; see also Rasmusen [1989]). It is recommended, however, not to announce in advance that the experiment will be
repeated, in order to guarantee a one-shot game situation in the first round.

Without repetitions, it takes about 15 minutes to conduct this classroom experiment; repetitions take about five minutes apiece. Thus, plenty of time is left for a discussion of the underlying incentive problem and how it can be overcome. Specifically, the possibility of defection introduces the question of institutional arrangements that help overcome moral hazard. Repetitions of the experiment using different information conditions - keeping the track records of participants secret or revealing them - are ideal to illustrate and motivate a discussion of third-party versus reputational enforcement in principal-agent interactions. Such issues are relevant to understand both modern theories of the firm as in Holmstrom and Tirole [1989], and the role of market forces in assuring contractual performance as in Klein and Leffler [1981].

III. LIKELY RESULTS

In our experience, the cooperative and collectively optimal outcome occurs about 30% of the time. Defection of agents, leading to the outcome (provide low quality, don't inspect), occurs about 50% of the time. The remaining 20% account for the two outcomes where the buyer inspects. (These percentages are "typical," but can vary significantly in any single experiment.)

As regards sellers, the percentage of successful cooperation, as well as the percentage of defection, are thus roughly the same in asymmetric dilemma of trust as in symmetric prisoner's dilemma games as in Marwell and Ames [1981], Carter and Irons [1991], and Frank et al. [1993]. The difference between the symmetric game problem and the asymmetric game problem is the choice set and the actual choices of the buyers (principals). Typically, buyers inspect only one out of five times.

It is important to remember that agent and principal face very different choices when our pay-off matrix is used. Specifically, agents have a (weakly) dominant strategy. In contrast, principals have the choice between a risky and a safe strategy. Given the relatively small payoffs, it is not surprising that a relatively small percentage of principal-students choose the safe strategy. (Their choice behavior reflects the fact that up to one dollar, students tend to be risk loving.) On the other hand, those agent-students that chose the (weakly) dominant strategy show behavior consistent with the behavior of populations in prisoner's dilemma game situations. Thus, the fact that students do not settle at the Nash equilibrium in this asymmetric situation is not surprising, and in fact is consistent with previous experimental results.

In class we compare these results with the result predicted by means of economic theory. We also ask how the outcome might have differed in other situations (for instance, guarantees that the participants were anonymous, that they would never deal with each other again, or would receive $1,000 payoffs, etc.)

We have conducted this experiment only with students in economics classes. It is possible that with other subject pools, one would get different results similar to those abundantly documented in two-sided prisoner's dilemma classroom experiments. Marwell and Ames [1981], in a famous article "Economists Free Ride, Does Anyone Else?" report that graduate students in economics are more likely to free-ride in public good provision experiments, that is, problems of the two-sided prisoner's dilemma variety, than other students.

Similarly, Carter and Irons [1991] investigate whether (aspiring) economists behave differently than other participants in simple ultimatum experiments. Employing undergraduate students, they confirm in their experiments that economics majors behave more in accordance with the predictions of game theory. They also investigate whether these results stem from self-selection or whether economists are successfully drilled in "the economic way of
thinking” during the course of their undergraduate studies. Their results are somewhat inconclusive.

Frank et al. [1993] continue this line of research and summarize, among other things, the major objections to both the Marwell and Ames and Carter and Irons studies. They address the question of whether the difference in behavior is the result of economics training through additional prisoner's dilemma experiments involving both economics and noneconomics majors. Frank et al., too, find that economics majors are more likely to act in accordance with the predictions of game theory. One of the intriguing results of their study is that gender also seems to play a significant role. New experiments that one of us has conducted recently, suggest that exposure to economics may be less important, if at all, than the gender composition of subject populations. Specifically, Ortmann and Tichy [1995] find significantly different cooperation rates of female and male students in prisoner's dilemma games. Controlling for these gender effects, exposure to economics does not seem to make a significant difference in subjects' choice behavior.

Finally, Yezer, Goldfarb, and Poppen [1996] report a "lost letter" experiment in which envelopes with ten dollars and an address were left in various upper-level classes in different fields. They find that economics students returned a significantly larger percentage of lost letters, exhibiting more cooperative behavior than other students, and conclude that studying economics does not discourage cooperation. Yezer et al. recommend that one should watch what economists do, not what they say or how they play.

This set of articles addresses philosophical, design, and curricular issues that, in our experience, have rarely failed to stimulate class discussion. All articles are easily accessible so that an undergraduate, possibly with some coaching, should be able to summarize and present them in class.

V. CONCLUDING COMMENTS

The advantages of experiments such as this one go far beyond the actual lesson of the experiment. The reasons are various; one is pedagogical. Experiments are hands-on; they allow students to experience economics. Experiments are thus one important way to address the documented "chilly" classroom climate reported by Siegfried et al. [1991a; 1991b] that seems in particular to affect female and/or minority students as noted by Catanese [1991], Shackelford [1992], and Siegfried and Scott [1994]. Experiments make students feel more welcome, and ultimately more receptive. A second advantage of experiments is that they can be presented as an alternative to econometric testing. As econometric work has come under increasing fire, alternatives look more desirable. A third advantage is that experiments demonstrate to students subtleties that would otherwise be missed. Specifically, experiments demonstrate the importance of institutional arrangements and how even slight changes in institutional design can lead to significant changes in outcomes, as in Davis and Holt [1993]. We find all these reasons to use experiments in the classroom persuasive as long as the teacher is explicit about the limitations of experimental demonstrations such as the one suggested here.

As a final comment it is important to note that the use of classroom experiments does not necessarily have to imply a lowering of standards of rigor, a point discussed in Ortmann and Scroggins [1995]. As we stated at the beginning, classroom experiments, if used judiciously, can motivate the students and introduce them to far more sophisticated economic ideas than the more routine presentation. We believe that this simple experiment is a useful springboard for discussion of a wide range of issues that belong in a microeconomics class.


2. An alternative parameterization that one also can find in the literature (see Friedman [1991]), is (hq, don't) = (1,2), (hq, i) = (0,1), (lq, don't) = (2,-1), and
(lq, i) = (-1, 0). If one uses it one must introduce mixed-strategy Nash equilibria, which is a bit difficult for introductory students. We chose our simpler parameterization because we felt that a discussion of mixed-strategy equilibria distracts from the moral hazard aspects of the principal-agent interaction. On the other hand, nothing about our set-up depends on our parameterization, and those who would like to introduce mixed-strategy equilibria will find that extension straightforward. See Kreps [1990a] and Binmore [1992] for excellent discussions of mixed-strategy equilibria.

3. Another parameterization was suggested by a referee. The referee suggested that (hq, don't) = (2,2), (hq, i) = (lq, i) = (1,1), and (lq, don't) = (3, -2). The parameterization is based on the following assumptions. The value of a high quality repair to the buyer is 6. The value of a low quality repair is 2. The price of the repair is 4. Thus, if he does not inspect, the buyer's net benefit of a high quality repair is 2, and of a low quality repair is -2. The repair person's cost of a high quality repair is 2 and the cost of a low quality repair is 1, giving her a net benefit of 2 for a high quality repair and of 3 for a low quality repair if the buyer does not inspect. The costs of inspection are, as in our story, assumed to be I for the seller and 1 for the buyer, leaving them each with a payoff of 1. "The intuition here is that the repairer might try to slip by doing low-cost (1 unit of cost) repairs. Inspection forces the repairer to 'finish the job' for an extra 1 unit. The first unit is not wasted; the job was just not complete."

Note that the referee's parameterization maintains the game-theoretic prediction of our parameterization. However, it changes the ordering of the cells - the joint payoff for (low quality, don't inspect) is now less than the joint payoffs in the cells of the right column. This weakens the moral hazard aspect of the experiment - instead of possibly doubling the payoff, the agent's potential gain from giving in to temptation is only 50%, making cheating/shirking behavior on the part of the agent less likely.

4. The above discussion is technically relevant only for finitely repeated games. It is a well-known theorem in game theory that any finitely repeated game has the same formal result as a one-shot game under certain assumptions. The reasoning has to do with backward induction. The agent can reason that in the last game there is no reason for developing trust, so she can assure that the other will choose the no-trust solution. But if she therefore provides a low quality repair in the last game, then she will also find it advantageous to choose it in the next to last game, and so on until one has reasoned back to the initial game. This backward induction approach applies to all finite games under complete information in which players are individual payoff maximizers.

All these issues have been the subject of much experimental and theoretical work. In the realm of theory, Kreps, Milgrom, Roberts, and Wilson [1982] have shown that incomplete information can change the counterintuitive game-theoretic predictions for finitely repeated games. A similar result for indefinitely repeated games, both with complete and incomplete information, has long been known as a folk theorem; see Rasmusen [1989] for a discussion and further references. For a simple algebraic argument using a so-called trigger strategy, see Kreps [1990a, 66-72]; see also Rasmusen [1989, 96-9], for an extension of the argument to situations involving many sellers and buyers. Experimentally, subjects in finitely repeated games with complete information play the cooperative outcome more often than predicted by the standard game-theoretic solution; see Davis and Holt [1993] for a discussion and further references. We will return to some of these results in section III.

5. Other scenarios are possible. The game could describe employee-employer interaction, with the employee (the seller of labor) as the agent and the employer (the buyer of labor) as the principal as in Aron [1990] and Kreps [1990b]. Another application is that of an employee, now the principal, agreeing to work or not, and an employer, now the agent, living up to his promises or reneging on them as in Kreps [1990a, 65-72].

6. It is likely that scaling up the payoffs will lead to qualitatively different results. As documented by Smith and Walker [1993], increasing payoffs causes experimental results to converge towards game-theoretic predictions.
7. Some readers may be concerned that the results do not converge to the Nash equilibrium of the one-shot game. This is not of concern to us. In general, we do not believe that the success of an experiment is a function of students settling on the Nash equilibrium. However, the fact that an experiment does not settle on the Nash equilibrium suggests a number of interesting questions regarding the assumptions of game theory and issues of experimental design.

8. This claim is based on our experience; Fels's [1993] point about the need for more rigorous evaluation is well taken.

9. For example, typically class time is limited. In addition, experimental economics employs conventions and standards that are compromised by the very nature of the classroom. The setting, which involves students and their teachers, might lead to good subject behavior, i.e., students making the choices that they believe their teacher expects them to make. Also, it is a common practice in experiments to motivate subjects with substantial monetary payoffs. Typically, experiments are calibrated such that subjects can earn two to three times the minimum wage, thus making them prohibitively expensive for the classroom.

REFERENCES


Bishop, Jerry E. " 'All for One, One for All'? Don't Bet on It." Wall Street Journal, 4 December 1986, 37.


Catanese, Anthony V. "Faculty Role Models and Diversifying the Gender and Racial Mix of Undergraduate Economics Majors." Journal of Economic Education, Summer 1991, 276-84.


-----.


**Thomson Gale Document Number:** A56327393