E C O N O M I C S B U L L E T I N

Experimental Economics Research: Is there an alternative to having huge research budgets?

Amy Moore
Southern Utah University

Michael Taylor
Berkeley Mathemarketics Group

Abstract

The need to pay subjects to participate in experiments places a major financial burden on experimental economists. In this paper, we conduct dictator games and find that there is no difference in the way student subjects split money and extra-credit points, an encouraging result that suggests that giving course points could be a viable alternative to giving out cash in economic experiments.

The first author would like to thank her dissertation committee at the University of Pennsylvania for their encouragement and ideas. Any remaining errors are ours.
Citation: Moore, Amy and Michael Taylor, (2007) "Experimental Economics Research: Is there an alternative to having huge research budgets?." *Economics Bulletin*, Vol. 3, No. 4 pp. 1-6
Submitted: December 11, 2006. **Accepted:** January 20, 2007.
URL: http://economicsbulletin.vanderbilt.edu/2007/volume3/EB-06C90006A.pdf

1. Introduction

Rewarding experimental subjects with cash in order to study their actions in controlled situations has long been a tradition in experimental economics (Smith 1976). However, the need to pay subjects to participate in experiments places a major financial burden on researchers, often limiting sample sizes and sometimes even eliminating entire projects for want of an adequate research budget. In this short paper, we explore whether an alternative way of rewarding subjects produces similar results in an experimental game setting, and find that subjects treat extra-credit points for a course grade similar to monetary incentives, a result which suggests that giving course points could be a viable alternative to giving out cash in certain experimental situations.

Much research has been conducted on the effect of financial incentives on performance in experimental tasks, and Camerer and Hogarth (1999) provide an excellent review of that literature. Our focus in this paper is not on the effect of monetary incentives per se, but on evaluating whether subjects' response to earning class points is similar to earning cash.

Brown Kruse and Thompson (2001) compared the use of class points with monetary rewards and found that subjects responded differently to the two reward mediums. Their study was designed to elicit the value of a risk mitigation measure in a low-probability high-consequence scenario. By contrast, in our study using the experimental dictator game, we found that student subjects treated class points no differently from cash payouts, with the average amount of money they decided to keep for themselves being statistically no different from the average number of points they kept for themselves. This finding is encouraging since it suggests that at least in the case of simple experimental games, it may be possible to alleviate the financial burden of conducting research studies through the use of grade point incentives.

2. Method

The study was conducted amongst students enrolled in a required course at a major research university in the United States. The exercise was conducted at the beginning of the semester in order to maintain the same high salience for class points for all students[1]. Students participated in the exercise in class as part of the usual class routine at the beginning of the class session. The instructor was assisted by a research assistant who distributed and collected the data collection instruments from the subjects and disbursed the cash payments at the end of the experiment. The entire exercise took about 20 minutes to complete.

[1] If this study had been conducted at the end of the semester, extra-credit points would have been non-salient for some students such as the ones already earning a high A in the course.

Subjects played the well-known "dictator game" twice; once with money and once with extra-credit points that counted towards their course grade[2]. Since the order in which they played the two dictator games can have an impact on their decisions, the order of play was randomized amongst the subjects. About half the subjects first played the dictator game in which they had to split 10 extra-credit points[3], and then played the game in which they had to split $10; the rest of the subjects played first with money and then with points.

We chose the dictator game for its simplicity. We wanted a game that was easy for subjects to understand since these were naive subjects and this would be the students' first exposure to playing economics games. Also, since subjects' decisions are not predicated on an opponent's decisions, the dictator game is ideally suited for comparing the direct effect of two alternative incentive schemes. By using the dictator game, we were able to implicitly establish a comparative equivalence between the money (10 dollars) and the points (10 extra-credit points) that each student had to split in the two games (s)he played.

The instruction sheet made it clear to the subjects that they would be splitting the money (or the points) with another student just like them. A single-blind payoff protocol was used, i.e. students participating in the exercise would never learn the decisions made by any specific subject. The specific decisions made by each student were thus private information, known only to the experimenters. For both dictator games, subjects completed a short quiz meant to gauge understanding of the game and its instructions before starting the actual game. We made it a point to "flash the money" in front of the students before the start of the exercise in order to make the monetary reward salient, and thereby also increase the credibility of the points reward.

3. Analysis

The basic question we would like to answer is – do subjects exhibit the same decision pattern when they split points as compared to when they split money in an experimental game? If there is no difference, it would suggest that giving extra-credit

[2] The dictator game was developed as a modification of the ultimatum game (Davis and Holt 1993) and involves a subject deciding on a voluntary contribution to a second subject. In the ultimatum game, the second subject has ultimatum power and can either accept of reject the division proposed, with zero payoffs to both subjects if s/he rejects the division. In the dictator game, the first subject has the dictator power to unilaterally decide on the split; the second subject does not get to accept or reject the offer as in the ultimatum game. The simplicity of the dictator game makes it an attractive one to use. As in the standard dictator game, subjects in our study were told that they had received $10 (or 10 extra-credit points). They could choose to keep all the money for themselves, or could decide to give a portion of it to their unknown co-player. Complete instructions that were used are available from the authors on request.

[3] The grades for the course were assigned on a curve. However, the grading included a provision for extra-credit points that would be added to a student's grade *after* the curving. The points earned in the exercise were thus truly points for extra credit.

points in a course may be a viable alternative to giving monetary rewards while conducting simple economic experiments.

Before we discuss the findings from the study, we would like to briefly dwell on statistical power (see Cohen 1988 for a comprehensive treatment of the topic) and its implications for the sample size in this study. When designing this study, we wanted to ensure that we had a sample size that gave us adequate statistical power since the nature of the study meant that failing to reject the null hypothesis of no significant difference was the desirable outcome. Statistical power, which is defined as the probability that a false null hypothesis will be correctly rejected, therefore needed to be high. Consulting Cohen's power tables for the hypothesis test that we would conduct, and assuming a "medium" effect size (Cohen's d=0.5) told us that a sample size of 50 would provide adequately high power of 0.93. The actual number of students participating in the study was 47; using Bissonnette's (2000) power calculator, the power for this study for a two-tailed paired-comparison t test at α=0.05 is an adequate 0.918.

The overall findings are shown in Table I. Recall that we had randomly assigned subjects to two different "order" conditions since the order in which they played could possibly have an effect on their decisions. Approximately half the subjects first decided on the amount of money they would like to keep for themselves and then decided on the number of points to keep[4].

In order to assess whether the order of play affected decisions, we conducted a repeated-measures analysis of variance since each subject provided two dependent measures: moneykept and pointskept. The two decisions made by each subject was the within-subjects factor in the repeated-measures ANOVA and the order of play was the between-subjects factor.

We found that there were no significant order effects. The test for the between-subjects effect of order had a calculated F value of $F_{1,45} = 0.012$, which was highly non-significant (attained level of significance p=0.91). In the test of within-subject contrasts, the interaction effect of the order of play and the within-subjects factor was also non-significant ($F_{1,45} = 1.52$, p=0.224), both of which rule out the possibility that order of play impacted subjects' decisions.

Since the repeated-measures ANOVA told us that the order of play was not a significant factor, we conducted a t test for a comparison of means on the entire dataset. A paired-samples t test was used since the same subject had made decisions both with points and with money. With the sample size of 47, the mean number of points kept by the subjects was 7.68 (out of 10) and the mean amount of money kept was $7.66 (out of $10). The calculated value of the t statistic (t_{46}) for the difference between the paired means of moneykept vs pointskept was -0.082 with an attained 2-tailed level of

[4] We will call this condition "ys" – the y from "money" appearing before the s from "points," while the other condition is referred to as "sy."

significance of 0.935, thereby failing to reject the null hypothesis of no difference. Thus, the two means are not significantly different from each other.

We also conducted separate paired-samples t tests for each "order" group. The 23 subjects who played in the "ys" condition kept a mean of 7.87 points for themselves, and kept 7.52 dollars for themselves on average. The difference between the two means was not significantly different from zero (t_{22} = -0.879, p=0.389). In the "sy" condition, there were 24 subjects, with the mean values for points and money being 7.50 and $7.79 respectively. Here too, as we would expect from the results of the repeated-measures ANOVA, the paired-samples t test failed to reject the null hypothesis of no significant difference between the two means (t_{23} = 0.864, p=0.397)[5].

4. Conclusion

All the statistical tests outlined above suggest that there was no significant difference in the way student subjects made decisions in dictator games involving a decision to split money and points. In other words, subjects' decisions with regard to class points were no different from the decisions that they made with monetary rewards. Although we did this study in the limited context of a simple experimental game, we find the results encouraging since they suggest a way to alleviate the monetary burden on researchers working in the area of experimental economics.

Some areas of future research merit mention. First, although this study tells us that decisions made with points as rewards are not different from decisions made with money as the reward, we do not yet know the "conversion" rate between the two currencies. In other words, we do not know how many points researchers need to give in order to motivate a student to put in the same effort as she would put in for a payment of x amount of dollars. Second, we do not know if grades exhibit the same kind of non-monotonic relationship as financial incentives do (see for example Gneezy and Rustichini 2000). The present study was merely a first step in exploring if there are possibly other ways to incentivize students for participation in experimental economics research, and it is heartening to note that awarding class points does seem to be a solution.

[5] Since the underlying distribution may not be normal, we also used the non-parametric Wilcoxon signed ranks test to compare the means for our within-subjects design, and got the same results. The Wilcoxon test on the entire sample suggested that the two means were not significantly different (Z = -0.213, p=0.83). The means for the ys subgroup and the sy subgroup were also not significantly different (Z = -1.035, p=0.3 and Z = -0.765, p=0.44 respectively).

References

Bissonnett, Victor L. (2000). "Statistical Applets," http://fsweb.berry.edu/academic/education/vbissonnette/applets/applets.html

Brown Kruse, Jamie, and Thompson, Mark A. (2001). "A Comparison of Salient Rewards in Experiments: Money and Class Points," *Economics Letters*, December 2001, 74, 1, 113-17.

Camerer, Colin F., and Hogarth, Robin M. (1999). "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," *Journal of Risk and Uncertainty*, 19, 7-42.

Cohen, Jacob. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, N.J.:Lawrence Erlbaum Associates.

Davis, Douglas D., and Holt, Charles A. (1993) *Experimental economics*. Princeton, N.J.: Princeton University Press.

Gneezy, Uri, and Rustichini Aldo. (2000). "Pay Enough or Don't Pay At All," *The Quarterly Journal of Economics*, August 2000, 115, 3, 791-810.

Smith, Vernon L. (1976). "Experimental Economics: Induced Value Theory," *American Economic Review*, May 1976, 66, 2, 274-279.

Table I

	N	moneykept		pointskept	
		Mean	Std Dev	Mean	Std Dev
All cases	47	7.66 [a]	1.833	7.68 [a]	1.721
order= ys	23	7.52 [b]	1.880	7.87 [b]	1.914
order= sy	24	7.79 [c]	1.817	7.50 [c]	1.532

[a,b,c] Means marked with the same letter were compared using a paired-samples t test, and are not significantly different from each other.