2009

# Game Theory for Playing Games: Sophistication in a Negative-Externality Experiment

JM Spraggon

RJ Oxoby

# GAME THEORY FOR PLAYING GAMES: SOPHISTICATION IN A NEGATIVE-EXTERNALITY EXPERIMENT

JOHN M. SPRAGGON and ROBERT J. OXOBY*

*We explore the extent to which the lack of Nash payoff maximization in experimental games is attributable to the "sophistication" of participants (i.e., their understanding of strategic decision making and profit-maximizing decisions). To this end, we compare the behaviors of sophisticated participants (i.e., those who have been exposed to the concepts of game theory) against those of a more standard subject pool in a moral hazard environment. Results suggest that sophisticated subjects are significantly more likely to adopt strategies predicted by standard theory and arrive at a Nash equilibrium.* (*JEL* C72, C91, C92, D63, D64)

## I. INTRODUCTION

A significant body of literature addresses the behavior of individuals in experimental games and how this behavior often deviates from theoretical predictions. Specifically, this literature raises a concern with the lack of observed behaviors supporting theoretical (Nash) predictions. For example, voluntary contribution experiments often yield significant deviations from the Nash predictions, and participants' behaviors typically respond to experimental treatments that have no effect on the Nash predictions (Holt and Laury, (2008); Laury and Holt, (2008)). Moreover, greater than Nash contributions continue even in relatively long treatments (50 rounds and more). This tendency to over-contribute has been attributed to reciprocal altruism and decision errors.[1] While these explanations may be correct, there may be reasons for these deviations which are supported by standard theory (e.g., Binmore 1999).

We conjecture that some of the observed differences from theoretical predictions may be due to inexperience with the concepts of maximizing behavior and strategic interactions. Simply put, if individuals do not know what constitutes optimal decision making, it should not be surprising that optimal decision making is not observed. As a result, individuals may rely on simple rules or heuristics to make decisions even though these rules may be suboptimal.

To test this conjecture, we conduct a series of moral hazard experiments with participants who vary in their familiarity with the concepts and tools of optimal and strategic decision making. Specifically, we compare the behaviors of a "sophisticated" subject pool with those of

*The authors thank Bill Morrison for his comments and suggestions and Emily Birtwell and Kendra McLeish for valuable research assistance. J.M.S. acknowledges financial support from the Social Science and Humanities Research Council of Canada. R.J.O. acknowledges financial support from the W. E. Upjohn Institute for Employment Research and the Institute for Advanced Policy Research (University of Calgary).

*Spraggon:* Associate Professor, Department of Resource Economics, University of Massachusetts Amherst, 80 Campus Center Way, Amherst, MA 01003. Phone 413-545-6651, Fax 413-545-5853, E-mail jmspragg@resecon.umass.edu

*Oxoby:* Associate Professor, Department of Economics, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada T2N 1N4. Phone 403-220-2586, Fax 403-282-5262, E-mail oxoby@ucalgary.ca

1. See Ledyard (1995) and Laury and Holt (2008) for reviews of this literature in public goods experiments. Similar arguments have been made regarding deviations from theoretical predictions in ultimatum and gift-giving games (Fehr and Fischbacher 2002) as well as a wide range of other games (Goeree and Holt 2001).

**ABBREVIATIONS**

MW: Mann-Whitney
KW: Kruskal-Wallis

doi:10.1111/j.1465-7295.2008.00146.x

a more standard pool of participants (i.e., undergraduate university students). In identifying sophisticated participants, we chose individuals who (at a minimum) had taken an undergraduate game theory course. Our rationale was that these students should be familiar not only with the idea of marginal analysis but also with the concepts of the Nash equilibrium and the identification of dominant strategies in games. Thus, our sophisticated subject pool can be thought of as having been trained in "payoff maximization."[2] Our interest lies in how these individuals behave relative to a more standard pool of "unsophisticated" participants. As such, our analysis is akin to that which attempts to induce behavior which is consistent with theoretical predictions such as Plott and Zeiler (2005), Cherry, Crocker, and Shogren (2003), and Charness, Frechette, and Kagel (2004).

In our environment, subjects choose decision numbers for which higher decision numbers correspond to higher individual payoffs and higher social costs, analogous to the emission of a pollutant that is costly to abate. Under the assumption that these decision numbers are private information, this is a classic moral hazard in groups problem similar to the worker effort problem in the labor literature (e.g., Holmstrom 1982) and the nonpoint source pollution in the environmental literature (e.g., Segerson 1988).

We use two instruments based on the family of exogenous targeting instruments suggested by Segerson (1988). These instruments involve an exogenously chosen target for aggregate (i.e., group) decision numbers, analogous to the aggregate environmental level of a nonpoint source pollutant. The two instruments we use create incentives for individuals to choose optimal decisions by providing penalties (a tax) or rewards (a subsidy) for implementing aggregate decisions greater than or less than the exogenous target.[3] Our Tax/Subsidy instrument involves a tax if the sum of individual decision numbers exceeds the target and a subsidy if the sum is below the target. Under this instrument, there is a unique, interior dominant strategy Nash equilibrium and a group (Pareto) optimal outcome. Our Tax instrument involves only the tax if the aggregate decision number of subjects in the group exceeds the target level. Under this instrument, there is a unique, interior Nash equilibrium (although it is not a dominant strategy). The differences between equilibria under each instrument allow us to discern how individuals' experience affects their ability to play equilibrium strategies and identify superior (i.e., Pareto-dominant) equilibria.

The environment investigated in this paper differs significantly from standard social dilemma experiments such as public goods, ultimatum, and gift exchange games. In these standard social dilemmas, there is a clear choice between self-interested and other-regarding plays. In our experiment, the instrument (either the Tax/Subsidy or the Tax) is designed so as to eliminate the social dilemma. As a result, both the self-interested and the other-regarding play outcomes are the same. This alignment of self-interested and other-regarding preferences should provide the theoretical prediction a better chance of being observed than in social dilemma experiments. That it does not (Cochard, Willinger, and Xepapadeas 2005; Poe et al. 2004; Spraggon 2002, 2004a, 2004b; Vossler et al. 2006) is a question that must be resolved before the behavior in social dilemma experiments can be fully understood.

Laury and Holt (2008) survey the literature on public goods games with interior Nash equilibria, concluding that moving to an interior Nash equilibrium does not result in decisions that are more consistent with the Nash predictions. They find that average decisions are typically between the Nash equilibrium and the midpoint of the decision space. In group moral hazard environments similar to that employed here, previous research has shown that under different conditions (e.g., market environments, communication), aggregate decisions are close to the Nash prediction, whereas individual-level decisions differ significantly from these predictions (Cochard, Willinger, and Xepapadeas 2005; Poe et al. 2004; Spraggon 2002, 2004a, 2004b; Vossler et al. 2006). In this paper, we show that the disparity between actual behavior and the Nash predictions can

2. An anonymous reviewer points out that these subjects may find game-theoretical principles more intuitive than our standard subject pool. We admit that this is a valid concern but feel that to the extent we are interested in the behavior of competitive agents (be it firms in the nonpoint source pollution problem or workers in the worker effort problem), presumably these agents are also selected from those who are better at optimization.

3. This type of mechanism is similar to the provision point public goods mechanisms analyzed by Bagnoli and Lipman (1989) and Bagnoli and McKee (1991).

be reconciled when the participants understand the concepts of the Nash equilibrium and dominant strategies.

Overall, we find that the behavior of sophisticated participants is much closer to the Nash predictions than that of unsophisticated participants. This is particularly true under the Tax instrument. In addition, we find the behaviors of sophisticated subjects to be much less volatile than those of unsophisticated subjects. This is true even when sophisticated subjects attempt to "signal" their desire to coordinate on the group optimal outcome. Thus, while training in economics may inhibit cooperativeness (Frank, Gilovich, and Regan 1993; Marwell and Ames 1981), this training may also improve decision making in environments embodying strategic behavior and information problems. Thus, our results are related to and build upon the work of Cherry, Crocker, and Shogren (2003) in which the rationality participants demonstrate as a result of market discipline spills over into a nonmarket setting. Our results complement this research, demonstrating that the rationality one is exposed to and practices in developing an understanding of game theory affect decision making in our negative-externality environment.

Our results suggest that the practical value of exogenous targeting instruments may be underestimated in a moral hazard environment where decision makers have experience with profit maximization (e.g., professionals in the field making decisions in nonpoint pollution and team production environments). On the other hand, the behavior of unsophisticated participants is much more volatile than that of sophisticated decision makers and is much less likely to be consistent with the Nash predictions. Perhaps surprisingly, we identify several "rules" or heuristics that appear to guide the decisions of unsophisticated individuals. That is, many of the decisions made by unsophisticated decision makers converge to focal points that are consistent with simple decision-making rules.

Our results provide a clear indication that an understanding of optimal decision making and Nash behavior (or a lack thereof) goes a long way in explaining the observed deviations from equilibrium predictions. While various preference specifications may account for the deviations, it is striking that these alternate motivations are essentially absent among our sophisticated decision makers, who differ from the unsophisticated pool not based on preferences but rather based on the understanding of basic economic theory.[4] When this knowledge is absent, individuals appear to use alternate (and naively reasonable) rules to motivate their decisions.

We continue as follows. Section II lays out the environment used in our experiments and describes the two instruments participants faced. In Section III, we analyze the results of the experimental data at the aggregate level, by participant type, and at the individual decision level. We find that sophisticated decision makers (i.e., those with an understanding of game theory and profit maximization) are more likely to make decisions supporting the Nash predictions. Furthermore, the variability of individual decisions appears to be significantly muted in experiments with sophisticated participants. In Section IV, we discuss our results, casting our findings in light of the behavior of individuals in experiments and the practical use of exogenous targeting instruments.

## II. EXPERIMENTAL DESIGN

Moral hazard in groups is inherent in situations as varied as the workplace (Holmstrom 1982), insurance (Rothschild and Stiglitz 1976), and the environment (Segerson 1988). The moral hazard in this experiment is due to a regulator wanting to reduce a negative externality resulting from consumption. This is analogous to an environmental problem in which unabated pollution maximizes firm profits but reduces social welfare. In the worker effort and insurance problems, the regulator seeks to increase the positive externality associated with increased effort (the more effort exerted by the worker, the better off the firm).[5]

In our experiment, groups consist of four participants, two of whom choose decision numbers between 0 and 100 (medium-capacity

4. Similar points have been made regarding refinements in signaling games (Banks, Camerer, and Porter 1994; Brandts and Holt 1993).

5. See Park (2001), Willinger and Ziegelmeyer (1999), Sonnemans, Schram, and Offerman (1998), and Andreoni (1995) for empirical comparisons of positive- and negative-externality environments.

participants) and two between 0 and 125 (large-capacity participants).[6] Both types of participants face the same private payoff function

$$(1) \quad B_n(x_n) = 25 - 0.002(x_n^{\max} - x_n)^2,$$

where $x_n^{\max} = 100$ for medium-capacity participants and $x_n^{\max} = 125$ for large-capacity participants. These payoffs are described to the participants by way of a table, and private payoff is maximized when $x_n = x_n^{\max}$.[7]

The moral hazard aspect of the experiment is implemented through an external cost[8] proportional to the aggregate decision number $X = \sum_{n=1}^{4} x_n$ given by

$$(2) \quad D(X) = 0.3X.$$

In this environment, individual decisions $x_n$ are private information, while $X$ is observable. Thus, a Paretian regulator interested in efficient aggregate outcomes should use instruments based on the aggregate decision number $X$ via an exogenous target $X^*$ (as in Holmstrom 1982; Segerson 1988). Given an aggregate decision number $X$ and an exogenous target $X^*$, each individual pays the tax (if $X > X^*$) or receives the benefit (if $X \leq X^*$) $T_n(X)$ given by

$$(3) \quad T_n(X) = \begin{cases} t_n(X - X^*) + \tau_n & \text{if } X > X^* \\ s_n(X - X^*) - \beta_n & \text{if } X \leq X^* \end{cases}.$$

For our experiment, we chose $X^* = 150$ and consider two instruments: a Tax/Subsidy instrument in which $t_n = s_n = 0.3$, $\tau_n = \beta_n = 0$ and a Tax instrument in which $t_n = 0.3$, $s_n = \tau_n = \beta_n = 0$. Thus, under the Tax/Subsidy instrument, an individual's private payoff is given by

$$\pi_n = 25 - 0.002(x_n^{\max} - x_n)^2 - 0.3(X - 150), \quad (4)$$

while under the Tax instrument, an individual's private payoff is given by

$$(5) \quad \pi_n = 25 - 0.002(x_n^{\max} - x_n)^2 - \begin{cases} 0.3(X - 150) & \text{if } X > 150 \\ 0 & \text{if } X \leq 150 \end{cases}.$$

We consider the Nash equilibria under each of these instruments. Under the Tax/Subsidy instrument, for any $X$, an individual's best response $x_n^*$ is

$$(6) \quad x_n^* = x_n^{\max} - 75.$$

This is also the solution for the Tax instrument if participants believe that $X \geq 150$. However, if subjects believe that $X < 150$, then their payoff-maximizing strategy is given by

$$(7) \quad x_n = \min(x_n^{\max}, 150 - X_{-n}).$$

where $X_{-n} = \sum_{j \neq n} x_j$. Whereas the Tax/Subsidy instrument results in a clear dominant strategy independent of the decisions of others, there is no such strategy under the Tax instrument. There is also a second (Pareto superior from the point of view of members of the group) optimum under the Tax/Subsidy instrument: if all participants choose $x_n = 0$, the payoff to the group is maximized.

The above analysis is based on the assumption that participants maximize their monetary payoff. We also consider the possibility that our subjects may be boundedly rational.[9] In the environment presented here, we are particularly interested in theories involving rules of thumb (Bagnoli and Lipman 1989; Hackett, Schlager, and Walker 1994; Rapoport and Suleiman 1993). For example, a simple decision rule in this environment is choosing a decision number equal to the target divided by the number of participants. Such a rule provides a simple way for the group to avoid paying a fine. Similarly, participants may make decisions based on focal points (e.g., the midpoint of their decision space).

## III. RESULTS

In this section, we present our experimental results. We first consider whether or not the

6. This environment is based on the moral hazard in group experiments conducted by Spraggon (2002).

7. Instructions and the payoff table are provided on the lead author's Web site: http://www.umass.edu/resec/faculty/spraggon/index.shtml.

8. Natural examples of this cost are pollution generated from the individual production decisions of individual firms or individuals in work teams free riding on the efforts of others in the team.

9. We could also consider the possibility that subjects are maximizing utility functions, which include variables other than their own payoff. In a similar environment, Spraggon (2004b) suggests that alternate preferences are not appropriate.

sophisticated groups are more consistent with the Nash prediction than the unsophisticated groups at the aggregate level. We then look at the decisions over participant type (medium or large capacity) and at the individual level.

The data were collected from eight sessions, each consisting of two groups of four subjects, conducted at our universities. Participants were recruited from economics courses and classified as sophisticated if they had taken an undergraduate game theory course.[10] Each group consisted of either all sophisticated or all unsophisticated individuals, and each experiment consisted of 25 decision-making periods under either the Tax instrument or the Tax/Subsidy instrument.[11] The decision was not presented as a maximization problem. Subjects were given a tabular version of Equation (1) and both verbal and mathematical descriptions of the group payoff function (Equation 4 or 5). The software used for the sessions provided the subjects with a calculator, allowing them to determine their payoff from different combinations of their decisions and the decisions of the others in the group. Sessions took approximately 90 min, and average earnings varied between $10 and $25 Canadian.

### *A. Analysis at the Aggregate Level*

Previous experiments by Spraggon (2002, 2004b) lead us to believe that the instruments will be effective in inducing groups to the Nash equilibrium at the aggregate level (i.e., $X = X^* = 150$). Here, while we expect variability in aggregate decision numbers, this variability should be significantly lower within groups of sophisticated participants. Indeed, this is demonstrated in Table 1, which presents the aggregate decision number $X$ by session. As expected, means are closest to $X^* = 150$ and less variable for groups of sophisticated participants under both instruments. The median aggregate decision numbers (Table 2) show this even more sharply. Note that only the aggregate decisions for the unsophisticated subjects under the Tax instrument is significantly different from the target of 150.

**TABLE 1**
Mean Aggregate Decision Numbers ($X$) by Treatment.

| Instrument | Unsophisticated | Sophisticated | Total |
|---|---|---|---|
| Tax | | | |
| Mean[a] | 209.96* | 153.07 | 175.41 |
| SE | 7.43 | 2.24 | 10.77 |
| *n*[b] | 3 | 5 | 8 |
| Tax/Subsidy | | | |
| Mean[a] | 181.37 | 152.13 | 164.66 |
| SE | 17.34 | 4.35 | 9.13 |
| *n*[b] | 3 | 4 | 7 |
| Total | | | |
| Mean[a] | 195.67 | 152.65 | 169.86 |
| SE | 10.59 | 2.14 | 7.02 |
| *n*[b] | 6 | 9 | 15 |

*Note:* SE: standard error.
[a]Mean of the mean aggregate decision number for each treatment over the number of sessions.
[b]Number of observations.
*indicates that the mean is significantly different from the target of 150 at the 5% level.

Analysis of variance on the aggregate data (Table 3) suggests that both participant type (sophisticated or unsophisticated) and instrument are significant (at the 10% level). The difference between sophisticated and unsophisticated groups is confirmed for the Tax instrument by the Mann-Whitney *U*-test (hereafter MW) and Kruskal-Wallis $\chi^2$ test (hereafter KW) (for the difference between sophisticated and unsophisticated, $p = .0253$ for the Tax and $p = .157$ for the Tax/Subsidy instrument [the $p$ values are the same for both tests]).[12] This suggests that sophisticated participants made, generally speaking, choices that were more consistent with the Nash predictions. Moreover, choices by these participants are much less variable than those by unsophisticated participants. The MW and KW tests do not indicate any significant differences across treatments for either unsophisticated ($p = .2752$) or sophisticated ($p = .6242$) subjects. Thus, we conclude that at the

10. In general, subjects may have been familiar with each other having taken courses together. We are not concerned that implicit cooperation is an issue here as no groups were able to coordinate on the group optimal outcome of all subjects choosing zero.

11. The experimental sessions were conducted by research assistants and not professors with whom participants may have had contact in their game theory courses.

12. We thank an anonymous reviewer for pointing out that the optimal decisions 25 and 50 are likely focal points for subjects, which makes finding significant differences between sophisticated and unsophisticated subjects less likely.

**TABLE 2**
Median Aggregate Decision Numbers by Treatment

| Treatment | Unsophisticated | Sophisticated | Total |
|---|---|---|---|
| Tax | | | |
| Median[a] | 205* | 150 | 153 |
| SE[b] | 9.49 | 1.20 | 11.46 |
| $n$[c] | 3 | 5 | 8 |
| Tax/Subsidy | | | |
| Median[a] | 199 | 150 | 150 |
| SE[b] | 21.20 | 1.25 | 10.61 |
| $n$[c] | 3 | 4 | 7 |
| Total | | | |
| Median[a] | 202 | 150 | 150 |
| SE[b] | 12.17 | 0.92 | 7.71 |
| $n$[c] | 6 | 9 | 15 |

[a]Median of the median aggregate decision number for each treatment over the number of sessions.

[b]Standard error based on the mean of the medians.

[c]Number of observations.

*indicates that the median is significantly different from the target of 150 at the 5% level.

aggregate level, sophistication matters for both Tax and Tax/Subsidy in the same way.

### *B. Analysis by Capacity*

Recall that subjects differed in their capacity (i.e., the size of their decision space). These differences in capacities permit us to analyze decisions to discern the different rules of thumb, which may have been used by unsophisticated subjects. Table 4 presents mean decision numbers by treatment, capacity, and five-period groupings. Note that decision making is very consistent with the Nash predictions for all except the unsophisticated medium-capacity subjects. Specifically, the means for the unsophisticated participants are much closer to 50 (the middle of their decision space) than 25 (the Nash equilibrium) under both the Tax and the Tax/Subsidy instruments. There are at least two potential explanations for the difference from the Nash prediction observed for the unsophisticated medium-capacity subjects. The first is confusion (random play), and the second is that they are maximizing their relative payoff by choosing higher numbers.[13]

For unsophisticated large-capacity subjects under the Tax instrument, Table 4 indicates that decisions are reasonably similar to the theoretical prediction (50). The decisions of sophisticated large-capacity subjects under the Tax instrument, however, are much more consistent with the theoretical prediction. This is evident by comparing the standard errors, medians, and modes between these two groups in Table 4. Both Levene's (1960) and Brown and Forsythe's (1974) tests for equality of variance suggest that the variances are significantly different ($p$ = .0000 for both of these tests). The distributions of individual decisions presented in Figure 1 also support this finding. These distributions are significantly different using the MW test ($p$ = .0000) and KW test ($p$ = .0001). For unsophisticated medium-capacity subjects under the Tax instrument, average, median, and modal decisions are much higher than the theoretical prediction of 25. This is not the case for sophisticated medium-capacity subjects under the Tax instrument whose decisions are very consistent with the theoretical prediction. Again, standard errors are significantly lower for sophisticated subjects in this treatment ($p$ = .000 for the Levene and Brown and Forsythe tests). Figure 2 shows the difference in the distributions of individual decisions between the unsophisticated and sophisticated subjects for this treatment. Again, the MW test ($p$ = .0000) and KW test ($p$ = .0001) confirm that these distributions are different.

As with the Tax, under the Tax/Subsidy instrument, decisions of the sophisticated subjects for both medium- and large-capacity subjects are completely consistent with the theoretical predictions. However, unsophisticated large-capacity subjects are less consistent with the theoretical prediction than they were under the Tax instrument. In both cases, the standard errors are significantly higher for the unsophisticated subjects (Tax/Subsidy large: $p$ = .0000 for the Levene test and $p$ = .0000 for the Brown and Forsythe test; Tax/Subsidy medium: $p$ = .0016 for the Levene test and $p$ = .001 for the Brown and Forsythe test). Figures 3 and 4 compare the distributions of individual decisions for medium- and large-capacity subjects. Again, the MW and KW tests both confirm that

13. Since everyone in the group pays the same fine, choosing higher numbers results in higher relative payoffs for the subjects choosing larger numbers.

**TABLE 3**
Analysis of Variance for Median Aggregate Results

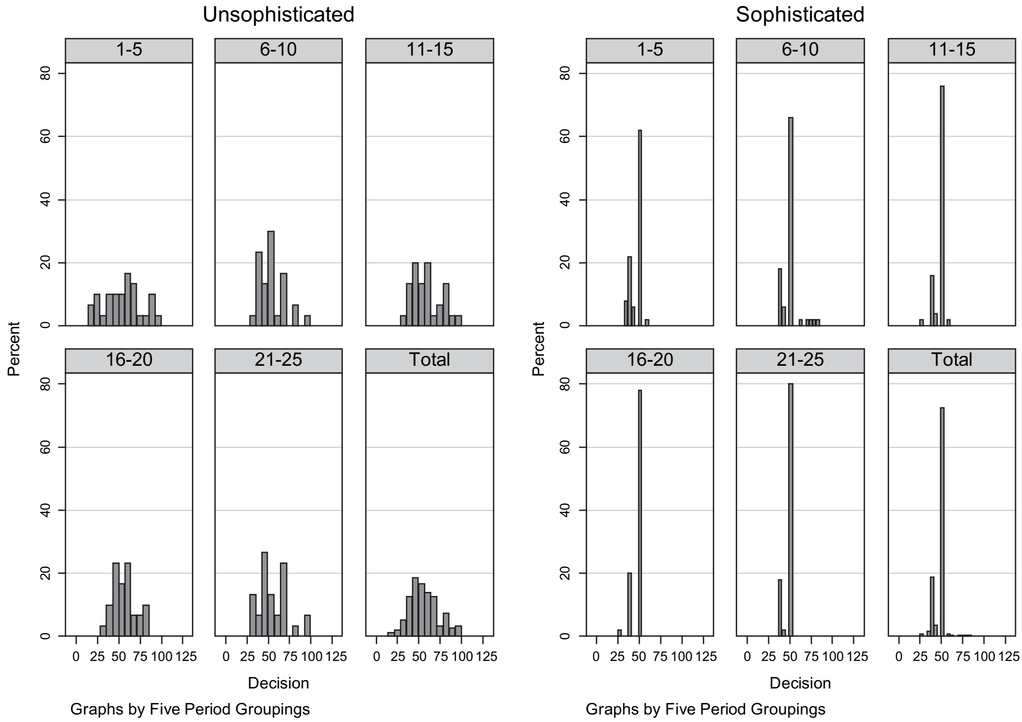
| Source | Partial SS | *df* | MS | *F* | Probability > *F* |
|---|---|---|---|---|---|
| Model | 7,888.30 | 3 | 2,629.43 | 11.74 | .0009 |
| Instrument | 780.84 | 1 | 780.30 | 3.49 | .0887 |
| Participant type | 6,643.53 | 1 | 6,643.53 | 29.67 | .0002 |
| Instrument × type | 684.38 | 1 | 684.38 | 3.06 | .1082 |
| Residual | 2,462.97 | 11 | 223.91 | | |
| Total | 10,351.26 | 14 | 739.38 | | |
| Number of observations | 15 | | | | |
| $R^2$ | 0.7621 | | | | |
| Root mean squared error | 14.96 | | | | |
| Adjusted $R^2$ | 0.6971 | | | | |

*Note:* Partial SS: partial sum of squares; *df*: degrees of freedom; MS: mean square; F: standard F-test.

**TABLE 4**
Individual Results by Five-Period Groups and Treatment

| | | Period | | | | | |
|---|---|---|---|---|---|---|---|
| Treatment | | 1–5 | 6–10 | 11–15 | 16–20 | 21–25 | Total |
| Large capacity, Tax, unsophisticated ($n$ = 30) | Mean | 55.2 | 53.37 | 59.03 | 56.33 | 56.37 | 56.06 |
| | SE | 4.15 | 2.85 | 3.07 | 2.54 | 3.23 | 1.43 |
| | Median | 57.5 | 50 | 57.5 | 50 | 51.5 | 50 |
| | Mode | 90 | 50 | 60 | 50 | 49.70 | 50 |
| Large capacity, Tax, sophisticated ($n$ = 50) | Mean | 45.96 | 50.06 | 47.52 | 47.24 | 47.72 | 47.7 |
| | SE | 0.87 | 1.36 | 0.816 | 0.793 | 0.660 | 0.422 |
| | Median | 50 | 50 | 50 | 50 | 50 | 50 |
| | Mode | 50 | 50 | 50 | 50 | 50 | 50 |
| Medium capacity, Tax, unsophisticated ($n$ = 30) | Mean | 46.97 | 47.43 | 48.67 | 48.83 | 52.7 | 48.92 |
| | SE | 3.25 | 3.42 | 3.17 | 2.25 | 2.56 | 1.32 |
| | Median | 50 | 46.5 | 48 | 50 | 50 | 50 |
| | Mode | 50 | 50 | 40 | 50 | 50 | 50 |
| Medium capacity, Tax, sophisticated ($n$ = 50) | Mean | 29.72 | 29.34 | 29.24 | 28.10 | 27.78 | 28.84 |
| | SE | 1.33 | 1.61 | 1.09 | 0.726 | 0.681 | 0.510 |
| | Median | 25 | 25 | 25 | 25 | 25 | 25 |
| | Mode | 25 | 25 | 25 | 25 | 25 | 25 |
| Large capacity, Tax/Subsidy, unsophisticated ($n$ = 30) | Mean | 53.5 | 53.57 | 47.7 | 47.87 | 44.93 | 49.51 |
| | SE | 3.77 | 3.37 | 3.26 | 2.43 | 1.87 | 1.36 |
| | Median | 51 | 51.5 | 41.5 | 45 | 45 | 46 |
| | Mode | 50 | 65 | 37.40 | 46 | 46 | 36.40 |
| Large capacity, Tax/Subsidy, sophisticated ($n$ = 50) | Mean | 49.96 | 49.33 | 47.95 | 49.93 | 52.65 | 49.96 |
| | SE | 1.52 | 0.857 | 0.622 | 1.11 | 2.61 | 0.680 |
| | Median | 50 | 50 | 50 | 50 | 50 | 50 |
| | Mode | 50 | 50 | 50 | 50 | 50 | 50 |
| Medium capacity, Tax/Subsidy, unsophisticated ($n$ = 30) | Mean | 46.37 | 39.4 | 34.03 | 41.90 | 44.17 | 41.17 |
| | SE | 4.48 | 2.92 | 3.80 | 3.87 | 4.62 | 1.79 |
| | Median | 42.5 | 40 | 40 | 39 | 38.5 | 40 |
| | Mode | 40.50 | 40 | 50 | 30 | 30.100 | 40 |
| Medium capacity, Tax/Subsidy, sophisticated ($n$ = 50) | Mean | 25.3 | 31.58 | 26.5 | 20.83 | 26.33 | 26.11 |
| | SE | 3.10 | 3.12 | 1.14 | 2.24 | 3.30 | 1.22 |
| | Median | 25 | 25 | 25 | 25 | 25 | 25 |
| | Mode | 25 | 25 | 25 | 25 | 25 | 25 |

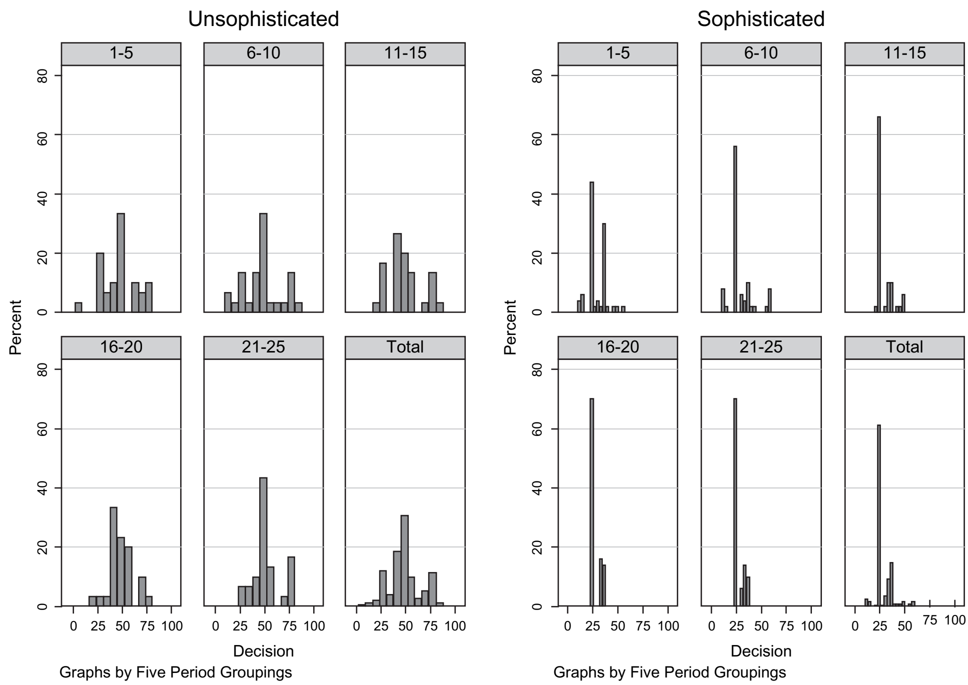
**FIGURE 1**
Distributions of Individual Decisions, Tax Instrument, Large-Capacity Subjects




the distributions of decisions for unsophisticated and sophisticated subjects are significantly different ($p$ = .0000 for the MW test for large-capacity subjects and $p$ = .0710 for medium-capacity subjects; $p$ = .0001 for the KW test for large-capacity subjects and $p$ = .0710 for medium-capacity subjects).

Table 4 and the distributions (Figures 1–4) do not suggest any dynamic adjustments. We use the nonparametric (MW and KW tests) and variance (the Levene and the Brown and Forsythe tests) comparison to compare the first and last five periods of each treatment to confirm this hypothesis. Under the Tax instrument, there is no significant difference between the first and last five periods for either unsophisticated or sophisticated medium-capacity or unsophisticated large-capacity subjects ($p > .16$ in all cases). For large-capacity sophisticated subjects, the difference in the distributions is close to significance ($p$ = .1019 for both the MW and the KW tests). Comparing the distributions in Figure 1 for these subjects, we see that decisions are a bit less random in the last five periods than they were in the first five periods for this group. Under the Tax/Subsidy, we observe a significant difference between the first and last five periods only for the unsophisticated large-capacity subjects ($p$ = .0132 for both the MW and the KW tests and $p > .54$ for all the other cases). Looking at the distributions in Figure 3, we again see that decisions are a bit less random by the last five periods in this case. In terms of variance, we observe significant differences between the variance of decisions in the first and last five periods for medium sophisticated and large sophisticated under the Tax instrument ($p < .052$ for these treatments and $p > .15$ for the other treatments) and large unsophisticated under the Tax/Subsidy instrument ($p < .01$ for this treatment and $p > .59$ for the other treatments). For the sophisticated subjects under the Tax instrument, Table 4 and Figures 1 and 2 indicate that decisions become more consistent with the theoretical prediction over time. Indeed, the unsophisticated large-capacity subjects are becoming as consistent

**FIGURE 2**
Distributions of Individual Decisions, Tax Instrument, Medium-Capacity Subjects

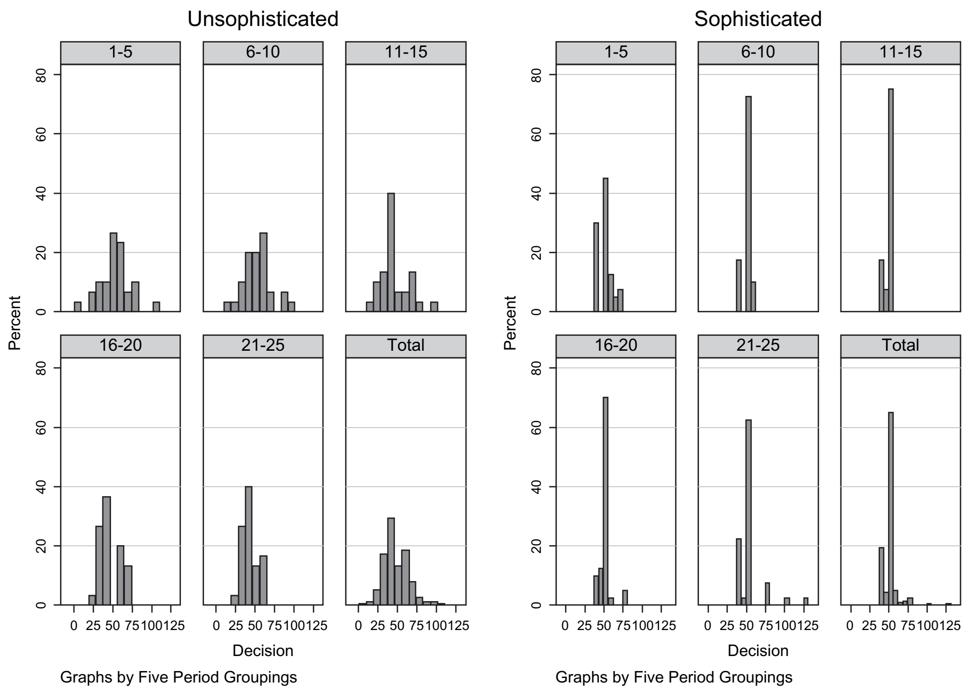



with the theoretical prediction by the last five periods as the sophisticated subjects are in the first five periods ($p$ = .0191 for both the MW and the KW tests).

Note that in addition to being more variable, the decisions of the unsophisticated participants have peaks at points conforming with the aforementioned simple rules. For large-capacity subjects, while we observe a peak at the Nash prediction of $x_n = 50$ under the Tax instrument, we also observe a large peak at the middle of the decision space ($x_n = 62.50$). With respect to these subjects under the Tax/Subsidy instrument, we observe two peaks, one to the right of the middle of the decision space and the other at the simple rule of dividing the target by the number of participants ($x_n$ = 37.50). The decisions of the medium-capacity unsophisticated subjects under both instruments are much more random. There is no peak at the Nash prediction under either of the instruments for this group, and the only large peak is in the middle of the decision space under the Tax instrument. This randomness may be attributable to the positioning of the Nash prediction relative to the alternate decision-making rule: for large-capacity subjects, the Nash prediction lies between the rule-of-thumb solution (target divided by the number of participants) and the simple heuristic solution (middle of decision space). This may help focus the decision making of large-capacity subjects. This is not the case for the medium-capacity subjects whose Nash prediction ($x_n = 25$) is below both the rule-of-thumb and simple heuristic solutions. As a result, medium-capacity subjects may exhibit greater "experimentation" in their decision making.

In contrast, the decisions of sophisticated subjects of both types and under each instrument are sharply centered on the Nash prediction ($x_n$ = 50 for large-capacity subjects and $x_n$ = 25 for medium-capacity subjects). We take this as strong evidence that sophisticated subjects are better able to understand how to make profit-maximizing decisions in this environment relative to unsophisticated participants. We find it telling the proportion of unsophisticated participants' decisions that are explained by our simple decision rules under the Tax instrument but surprising the

**FIGURE 3**
Distributions of Individual Decisions, Tax/Subsidy Instrument, Large-Capacity Subjects




degree of randomness under the Tax/Subsidy instrument.

## C. *Analysis by Participant*

We now consider the decisions of individual participants. Figures 5–8 present each subject's time series of decisions. With respect to unsophisticated subjects (Figures 5 and 6), we see significant volatility and little convergence to the Nash prediction. Indeed, under the Tax instrument, we observe only 2 of 12 subjects choosing their Nash decision numbers by the end of the 25 decision-making periods, and 3 of the 12 have median decisions that are equal to the Nash prediction. Similarly, only 3 of 12 subjects under the Tax/Subsidy instrument arrive at their Nash prediction, while none have medians that are equal to this value.

This stands in sharp contrast to the behavior of sophisticated subjects (Figures 7 and 8). Under the Tax instrument, the median decision of 14 of the 20 subjects is equal to the Nash prediction (with 6 subjects always choosing Nash and 4 others showing almost no deviation from this decision). Note that while Subjects 205 and 206 did not choose the Nash prediction, their behavior is optimal given that the other subjects in their group (207 and 208) chose slightly below their Nash predictions. Taken together, this yields 95% of sophisticated subjects under the Tax instrument whose decisions are consistent with the predictions of a Nash equilibrium. Under the Tax/Subsidy instrument, decisions are somewhat more volatile. However, the median decision of 12 of 16 subjects is equal to the Nash prediction. Some of the increased volatility (particularly that seen in Subjects 102, 901, and 1501) may be the result of trying to signal to other participants a willingness to move to the Pareto-superior outcome in which each individual chooses $x_n = 0$.

Again, we see a marked difference between the behavior of sophisticated and that of unsophisticated participants. It is particularly striking how many of the sophisticated participants immediately identify the dominant strategy and play this strategy consistently (or more or less consistently). It is clear that

**FIGURE 4**
Distributions of Individual Decisions, Tax/Subsidy Instrument, Medium-Capacity Subjects

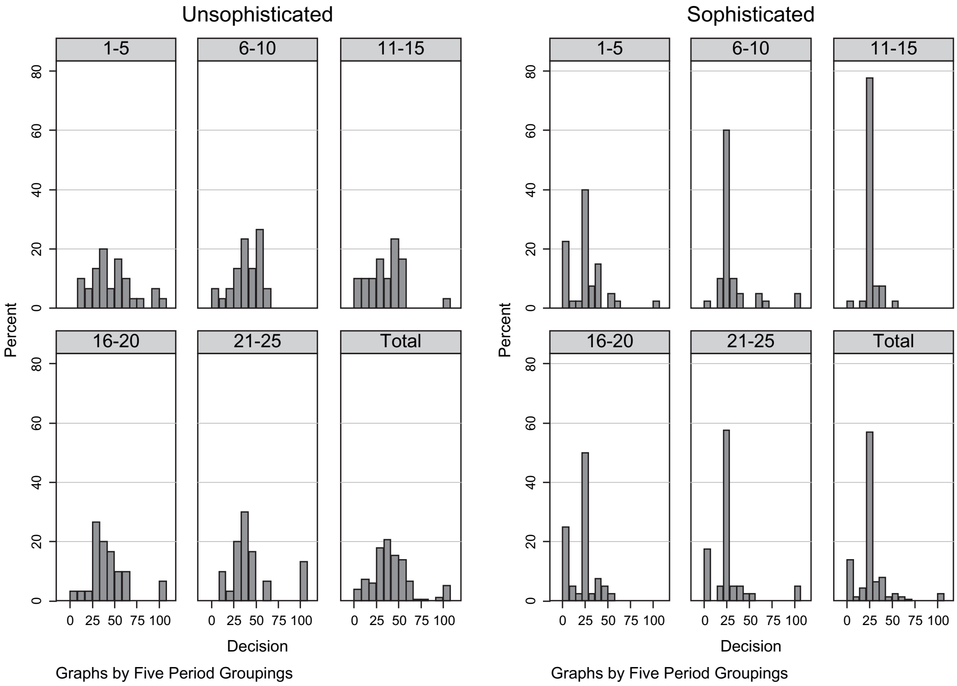



**FIGURE 5**
Individual Decisions, Unsophisticated Subjects, Tax Instrument

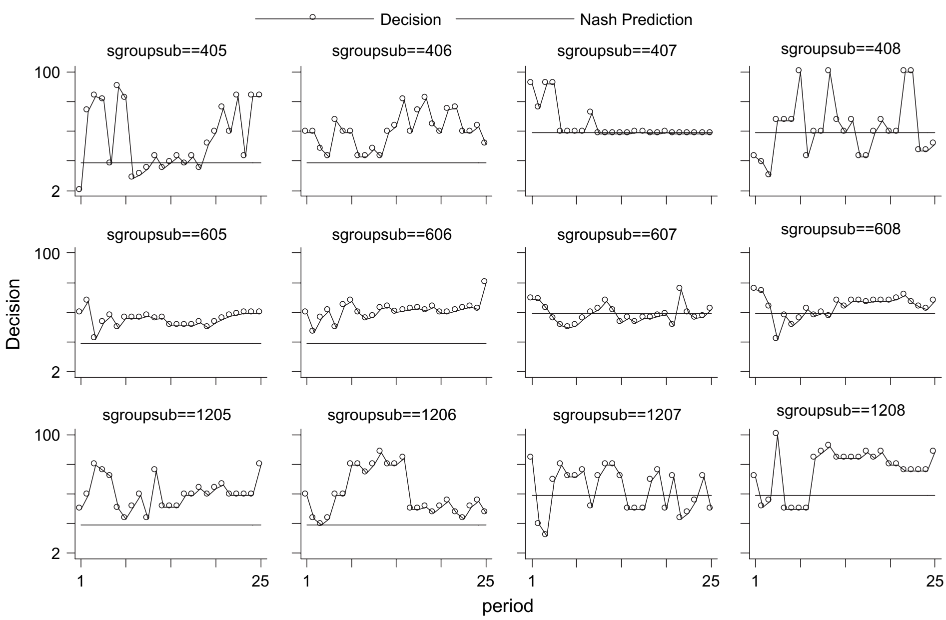


**FIGURE 6**
Individual Decisions, Unsophisticated Subjects, Tax/Subsidy Instrument

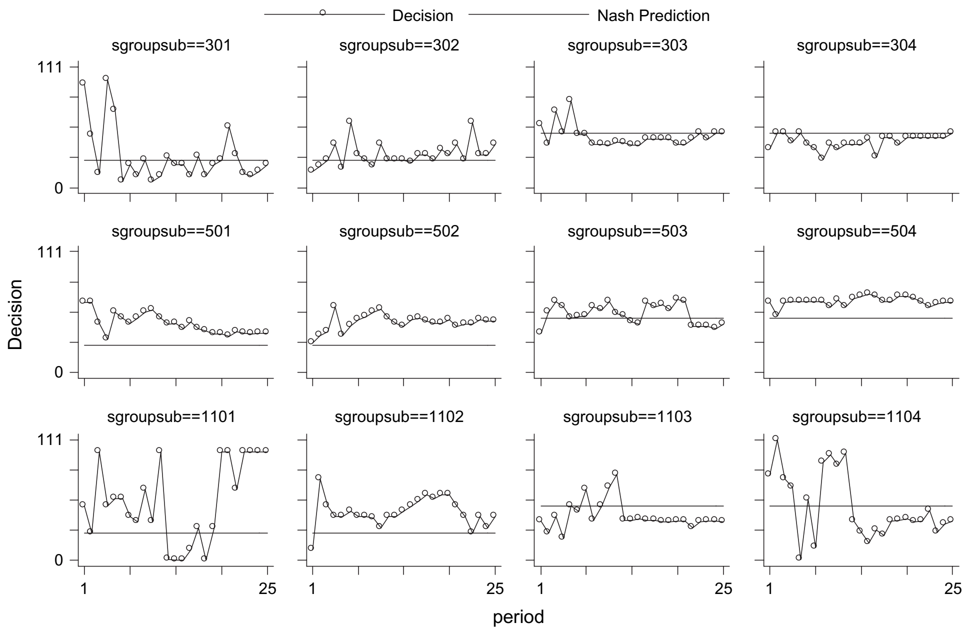



**FIGURE 7**
Individual Decisions, Sophisticated Subjects, Tax Instrument

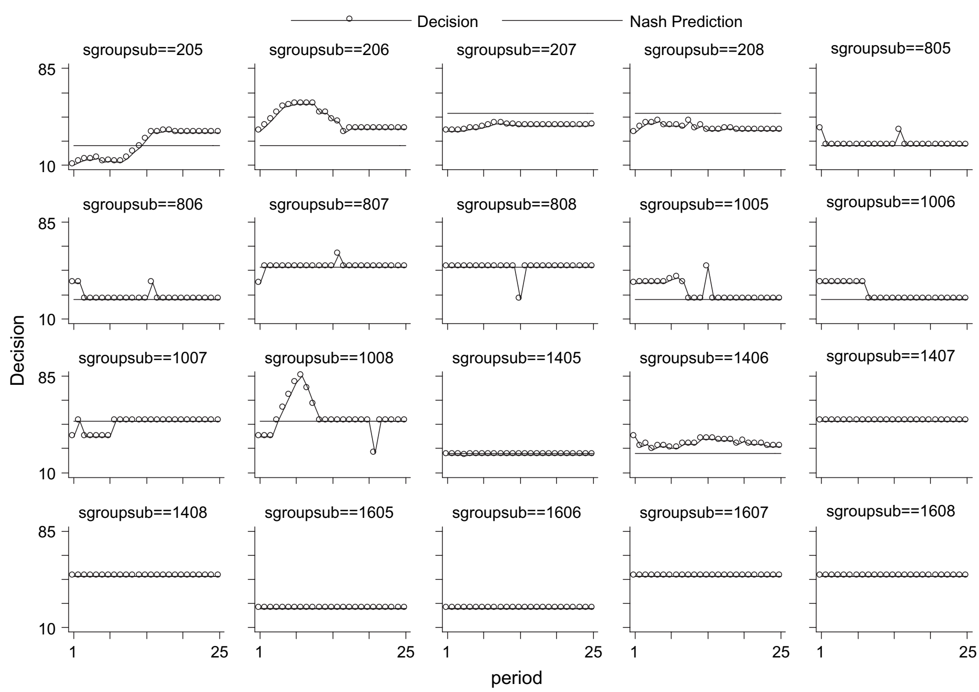
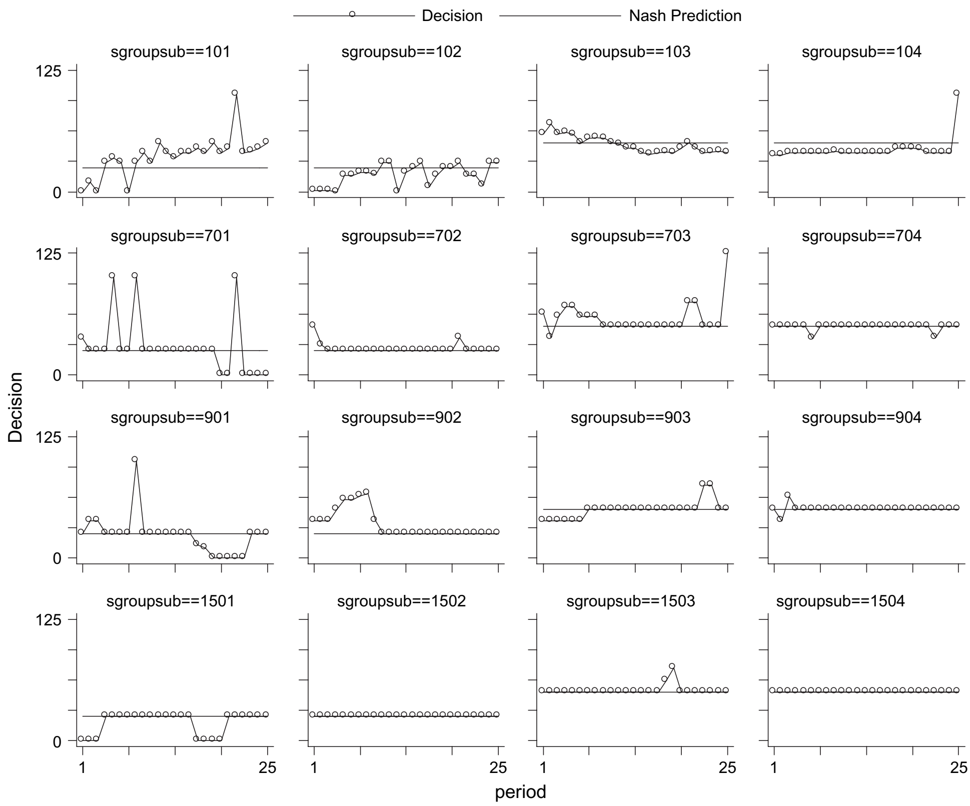


**FIGURE 8**
Individual Decisions, Sophisticated Subjects, Tax/Subsidy Instrument




these subjects understand the concept of profit maximization in a strategic environment and the idea of a dominant strategy. As such, it is perhaps unsurprising that their performance is so well predicted by a traditional Nash best-response strategy. On the other hand, the fact that the decisions of unsophisticated subjects stand in such sharp contrast to those of the sophisticated emphasizes the importance of the understanding of profit maximization when evaluating the performance of participants in an experimental game.

## IV. CONCLUSIONS

Our results demonstrate that when subjects understand optimization, their behavior is rationalizable and predicted by standard economic theory. In contrast, the behavior of subjects who are not trained in game theory (our unsophisticated subjects) displays evidence of the use of simple rules for making decisions. Having taken a course in game theory significantly reduces the decision cost of finding an optimal or a dominant strategy, suggesting that an extreme form of bounded rationality (one that is not necessarily based on Nash payoff maximization) would be more consistent with the behavior of the standard subject pool used for these types of experiments. This contention is consistent with Goeree and Holt (2001) who conclude that bounded rationality models based on initial beliefs coupled with experiments that elicit these beliefs is the most profitable approach to explaining individual-level decision making. Using subjects who are trained in game theory helps to control initial beliefs as not only do the individuals understand game theory, but also they know that the other people in their group understand game theory. This mitigates at least some of the strategic uncertainty found with the standard pool. Moreover (and perhaps thankfully), our results are consistent with those of Cherry, Crocker, and Shogren (2003) in that the rationality one acquires in taking a game theory course carries over into decision making in a negative-externality environment.

In our assessment, two implications derive from our results. Having course work in game theory had a significant effect on the behaviors of individuals. This is not surprising: if one does not know how to identify or what constitutes a dominant strategy, it is unlikely that such a strategy will be identified or chosen by a participant. It is not that the concept of dominance is not predictive but rather that participant inexperience in a relatively complex decision environment makes it highly unlikely that such a concept will be readily applied in decision making. Our results imply that in the many experiments where we fail to observe equilibria which support theoretical predictions, inexperience with the mechanics of optimization or strategic thinking may be to blame. Thus, the absence of behavior confirming theoretical conjectures may not be due to individuals having ulterior motives, decision errors, or unaccounted for arguments in their preferences but rather due to a basic naïveté with the types of decision making posited by rational choice theory.[14] More importantly, our results indicate that once people understand the tools of profit maximization and strategic decision making, they are relatively quick to implement and apply these concepts.[15]

The fact that individuals who know these concepts are able to implement ideas of profit maximization and play dominant strategies leads to a second implication of our results. The use of exogenous targeting instruments has often been criticized with respect to its application to environmental economics (Shortle and Horan 2001). Thus, while exogenous targeting instruments serve as a natural mechanism to cope with the problems of nonpoint pollution, they are rarely observed in practice. The relative dearth of these instruments may be attributable to concerns regarding how individuals behave when confronted with these types of incentives. Our results indicate that these instruments may work very well at implementing efficient allocations. In the economic environments where these instruments are most likely to be implemented, decision making is done by business people and entrepreneurs who, for the most part, are familiar with the idea of profit maximization, the functioning of markets, and strategic interactions.[16] With these sophisticated decision makers, the potential of these instruments to implement desired aggregate pollution levels in an efficient manner is greatly enhanced.

14. In a similar vein, Charness, Frechette, and Kagel (2004) find that behavior in gift exchange experiments is sensitive to the presentation of payoff tables, which more overtly indicate payoff-maximizing decisions.

15. Levitt and List (2007) make a related point regarding external validity and field experiments.

16. Potters and van Winden (2000) suggest that the decisions of professionals are more consistent with the Nash equilibrium than those of students in a lobbying game.

## REFERENCES


Andreoni, J. "Warm Glow versus Cold Prickle: The Effects of Positive and Negative Framing on Cooperation." *Quarterly Journal of Economics*, 1995, 110, 1–22.

Bagnoli, M., and B. Lipman. "Provision of Public Goods: Fully Implementing the Core through Voluntary Contributions." *Review of Economic Studies*, 1989, 56, 583–601.

Bagnoli, M., and M. McKee. "Voluntary Contribution Games: Efficient Private Provision of Public Goods." *Economic Inquiry*, 1991, 29, 351–66.

Banks, J., C. Camerer, and D. Porter. "An Experimental Analysis of Nash Refinements in Signaling Games." *Games and Economic Behavior*, 1994, 6, 1–31.

Binmore, K. "Why Experiment in Economics?" *Economic Journal*, 1999, 109, F16–F24.

Brandts, J., and C. A. Holt. "Adjustment Patterns and Equilibrium Selection in Experimental Signaling Games." *International Journal of Game Theory*, 1993, 22, 279–302.

Brown, M. B., and A. B. Forsythe. "Robust test for the equality of variances." *Journal of the American Statistical Association*, 1974, 69, 364–7.

Charness, G., G. R. Frechette, and J. H. Kagel. "How Robust Is Laboratory Gift-Exchange?" *Experimental Economics*, 2004, 7, 189–205.

Cherry, T. L., T. D. Crocker, and J. F. Shogren. "Rationality Spillovers." *Journal of Environmental Economics and Management*, 2003, 45, 63–84.

Cochard, F., M. Willinger, and A. Xepapadeas. "Efficiency of Nonpoint Source Pollution Instruments: An Experimental Study." *Environmental and Resource Economics*, 2005, 30, 393–422.

Fehr, E., and U. Fischbacher. "Why Social Preferences Matter—The Impact of Non-Selfish Motives on Competition, Cooperation, and Incentives." *Economic Journal*, 2002, 112, C1–C33.

Frank, R. H., T. Gilovich, and D. T. Regan. "Does Studying Economics Inhibit Cooperation?" *Journal of Economic Perspectives*, 1993, 7(2), 159–71.

Goeree, J. K., and C. A. Holt. "Ten Little Treasures of Game Theory, and Ten Intuitive Contradictions." *American Economic Review*, 2001, 91, 1402–22.

Hackett, S., E. Schlager, and J. Walker. "The Role of Communication in Resolving Commons Dilemmas: Experimental Evidence with Heterogeneous Appropriators." *Journal of Environmental Economics and Management*, 1994, 27, 99–126.

Holmstrom, B. "Moral Hazard in Teams." *Bell Journal of Economics*, 1982, 13, 324–40.

Holt, C. A., and S. K. Laury. "Theoretical Explanations of Treatment Effects in Voluntary Contributions Experiments," in *Handbook of Experimental Economic Results*, edited by C. Plott and V. Smith. New York: Elsevier Press, 2008.

Laury, S. K., and C. A. Holt. "Voluntary Provision of Public Goods: Experimental Results with Interior Nash Equilibria," in *Handbook of Experimental Economic Results*, edited by C. Plott and V. Smith. New York: Elsevier Press, 2008.

Ledyard, J. O. "Public Goods: A Survey of Experimental Research," in *Handbook of Experimental Economics*, edited by J. H. Kagel and A. E. Roth. Princeton, NJ: Princeton University Press, 1995.

Levene, H. "Robust tests for equality of variances," in *Contributions to Probability and Statistics: Essays in honor of Harlod Hotelling*, edited by I. Olkin, S. G. Ghurye, W. Hoeffding, et al. Menlo Park, CA: Stanford University Press, 1960, 278–92.

Levitt, S. D., and J. A. List. "Viewpoint: On the Generalizability of Lab Behaviour to the Field." *Canadian Journal of Economics*, 2007, 40, 347–70.

Marwell, G., and R. E. Ames. "Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods, IV." *Journal of Public Economics*, 1981, 15, 295–310.

Park, E.-S. "Warm-Glow versus Cold-Prickle: A Further Experimental Study of Framing Effects on Free-Riding." *Journal of Economic Behavior and Organization*, 2001, 43, 405–21.

Plott, C. R., and K. Zeiler. "The Willingness to Pay-Willingness to Accept Gap, the 'Endowment Effect', Subject Misconceptions, and Experimental Procedures for Eliciting Valuations." *American Economic Review*, 2005, 95, 530–45.

Poe, G. L., W. D. Schulze, K. Segerson, J. F. Sutter, and C. A. Vossler. "Exploring the Performance of Ambient-Based Policy Instruments When Nonpoint Source Polluters Can Cooperate." *American Journal of Agricultural Economics*, 2004, 86, 1203–10.

Potters, J., and F. van Winden. "Professionals and Students in a Lobbying Experiment: Professional Rules of Conduct and Subject Surrogacy." *Journal of Economic Behavior and Organization*, 2000, 43, 499–522.

Rapoport, A., and R. Suleiman. "Incremental Contribution in Step-Level Public Goods Games with Asymmetric Players." *Organizational Behavior and Human Decision Processes*, 1993, 55, 171–94.

Rothschild, M., and J. Stiglitz. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, 1976, 90, 629–50.

Segerson, K. "Uncertainty and Incentives for Nonpoint Pollution Control." *Journal of Environmental Economics and Management*, 1988, 15, 87–98.

Shortle, J. S., and R. D. Horan. "The Economics of Nonpoint Pollution Control." *Journal of Economic Surveys*, 2001, 15, 255–89.

Sonnemans, J., A. Schram, and T. Offerman. "Public Good Provision and Public Bad Prevention: The Effect of Framing." *Journal of Economic Behavior and Organization*, 1998, 34, 143–61.

Spraggon, J. "Exogenous Targeting Instruments as a Solution to Group Moral Hazard." *Journal of Public Economics*, 2002, 84, 427–56.

———. "Individual Decision Making in a Negative Externality Experiment." *Experimental Economics*, 2004a, 7, 249–269.

———. "Testing Ambient Pollution Instruments with Heterogeneous Agents." *Journal of Environmental Economics and Management*, 2004b, 48, 837–856.

Vossler, C. A., G. L. Poe, W. D. Schulze, and K. Segerson. "Communication and Incentive Mechanisms Based on Group Performance: An Experimental Study of Nonpoint Pollution Control." *Economic Inquiry*, 2006, 44, 599–613.

Willinger, M., and A. Ziegelmeyer. "Framing and Cooperation in Public Good Games: An Experiment with an Interior Solution." *Economics Letters*, 1999, 65, 323–28.