INCENTIVE-COMPATIBLE MECHANISMS FOR PURE PUBLIC GOODS: A SURVEY OF EXPERIMENTAL RESEARCH

YAN CHEN

School of Information, The University of Michigan, Ann Arbor, MI 48109-1092, USA
e-mail: yanchen@umich.edu

1. Introduction

The presence of public goods seriously challenges traditional or "natural" solutions for the allocation of private goods. Important policy questions, of whether we can rely on the market to provide optimal amounts of public goods such as air pollution, and how much we can rely on "natural" processes such as voluntary contribution to solve environmental problems, boil down to fundamental issues about human nature, i.e., about whether people are selfish or cooperative. There has been an extensive experimental literature that tried to answer this question and to evaluate the extent of the free-rider problem in environments with public goods. Ledyard (1995) concluded in his survey that

> "although inexperienced subjects can be led to provide large contributions in one-time decisions with the use of relevant discussions, one cannot rely on these approaches as a permanent organizing feature without expecting an eventual decline to self-interested behavior.[1] ... Since 90 percent of subjects seem to be responsive to private incentives, it is possible to create new mechanisms which focus that self-interest toward the group interest." (p. 173, "Public Goods: A Survey of Experimental Research." In: *The Handbook of Experimental Economics*, 1995.)

This article surveys experimental research on these "new mechanisms," i.e., incentive-compatible mechanisms for pure public goods.

1.1. Theoretical Results and Unresolved Issues

Hurwicz (1972) formally introduced the concept of incentive compatibility, which captures the forces for individual self-interested behavior. The theory of mechanism design treats incentive compatibility as a constraint on the choice of procedures used to make group allocation decisions in various economic contexts. The task of the mechanism

[1] Mark and Matthews (2000) consider the dynamic private provision of a public project. They show that under certain conditions perfect Bayesian equilibria exist that essentially complete the project.

Handbook of Experimental Economics Results, Volume 1

DOI: 10.1016/S1574-0722(07)00067-4

designer, therefore, is to find a mechanism such that the performance of the mechanism under an *assumed behavioral rule* is consistent with some normative performance criterion (e.g., Pareto-efficiency) given a class of environments. Experiments confront theory in these assumed behavioral rules.

"A fundamental, but generally unstated axiom of non-cooperative behavior is that if an individual has a dominant strategy available, he will use it" (Groves and Ledyard, 1987, p. 56) Therefore, theoretically it is desirable to design dominant strategy mechanisms, i.e., mechanisms which are non-manipulable. However, by now it is well known that it is impossible to design a mechanism for making collective allocation decisions, which is informationally decentralized, non-manipulable and Pareto optimal. This impossibility has been demonstrated in the work of Green and Laffont (1977), Hurwicz (1975), Roberts (1979), Walker (1980) and Mailath and Postlewaite (1990) in the context of resource allocation with public goods. The Vickrey–Clarke–Groves mechanism (Vickrey, 1961; Clarke, 1971; Groves, 1973; Groves and Loeb, 1975) admits dominant strategies but the allocation is not fully Pareto-efficient.

There are many "next-best" mechanisms which preserve Pareto optimality at the cost of non-manipulability, some of which preserve "some degree" of non-manipulability. Some mechanisms have been discovered which have the property that Nash equilibria are Pareto optimal. These can be found in the work of Groves and Ledyard (1977), Hurwicz (1979), Walker (1981), Tian (1989), Kim (1993), Peleg (1996), Falkinger (1996) and Chen (2002). Other implementation concepts include perfect Nash equilibrium (Bagnoli and Lipman, 1989), undominated Nash equilibrium (Jackson and Moulin, 1992), subgame perfect equilibrium (Varian, 1994b), strong equilibrium (Corchon and Wilkie, 1996), and the core (Kaneko, 1977), etc. Apart from the above non-Bayesian mechanisms, Ledyard and Palfrey (1994) propose a class of Bayesian Nash mechanisms for public goods provision.

To make any of these mechanisms operational and put it to use as an actual economic process that solves fundamental social problems, it is important to observe and evaluate the performance of the mechanism in the context of actual decision problems faced by real people with real incentives. These situations can be created and carefully controlled in a laboratory. When a mechanism is put to test in a laboratory, behavioral assumptions made in theory are most seriously challenged. More specifically,

1. Perfect vs Bounded Rationality: theory assumes that people are perfectly rational. As a result they can reach the equilibrium instantaneously through introspection. Since real people are boundedly rational, they need to learn by trial and error. This leads to an important aspect of mechanism design that has not received much attention: does a mechanism provide incentives for agents to learn?
2. Static vs Dynamic Games: since perfectly rational agents can reach equilibrium instantaneously, it is sufficient to restrict attention to static games. When a mechanism is implemented among boundedly rational agents, we expect the actual implementation to be a dynamic process, starting somewhere off the equilibrium path. This raises two questions:

(a) Can the learning dynamics lead to convergence to one of the equilibria promised by theory?
(b) What learning algorithms should be used to study the dynamic stability of a mechanism? This question can only be answered by estimating a rich repertoire of learning algorithms across a wide variety of experiments.

3. The Dominant Strategy Axiom: will agents use dominant strategies? If not, what other aspects might be important?
4. Learnability: what aspects of a mechanism might help agents to learn to play their Nash equilibrium strategies?
5. Refinement Criteria: do people learn to follow certain refinements of Nash equilibrium?

Despite the proliferation of theoretical literature on incentive-compatible mechanisms there have not been many experimental studies of these mechanisms. The existing experimental research on incentive-compatible mechanisms provides some data on the dynamic paths of these mechanisms when they are implemented among boundedly rational individuals. Some of these data have been used to investigate new theories on the dynamic stability of these mechanisms, incorporating bounded rationality and learning. The combination of theory and experimental results is likely to provide a fresh perspective on the mostly static implementation theory. They also raise many interesting questions, which call for further experimental as well as theoretical investigations.

1.2. Economic Environments in Experiments

Before reviewing the experimental results, we first introduce notation and the economic environment. Most of the experimental implementations of incentive-compatible mechanisms use a simple environment. Usually there is one private good x, one public good y, and $n \geqslant 3$ players, indexed by subscript i. Production technology for the public good exhibits constant returns to scale, i.e., the production function, $f(\cdot)$, is given by $y = f(x) = x/b$, for some $b > 0$. Preferences are largely restricted to the class of quasilinear preferences.[2] Let E represent the set of transitive, complete and convex individual preference orderings, $\geqslant_i$, and initial endowments, w_i^x. We formally define E^Q as follows.

DEFINITION 1. $E^Q = \{(\geqslant_i, w_i^x) \in E\ :\geqslant_i$ is representable by a C^2 utility function of the form $v_i(y) + x_i$ such that $Dv_i(y) > 0$ for all $y > 0$, and $w_i^x > 0\}$, where D^k is the kth order derivative.

The chapter is organized as follows. Section 2 reviews experimental studies of dominant strategy mechanisms. Section 3 reviews experiments on Nash-efficient mechanisms and introduces theoretical results on the convergence of these mechanisms. Section 4

[2] Smith (1980), Harstad and Marrese (1982) and Falkinger et al. (2000) are exceptions.

reviews experiments on mechanisms, which use refinements of Nash as implementation concepts. Section 5 reviews experiments on the Smith Auction. Section 6 concludes the chapter.

2. Dominant Strategy Mechanisms

When preferences are quasi-linear, the Vickrey–Clarke–Groves (VCG) mechanism is strategy-proof, where reporting one's preferences truthfully is always a dominant strategy. It has also been shown that any strategy-proof mechanism selecting an efficient public decision at every profile must be of this type (Green and Laffont, 1977). A special case of the Vickrey–Clarke–Groves mechanism is known as the pivotal mechanism. The pivotal mechanism has been tested in the field and laboratory by various groups of researchers.

Scherr and Babb (1975) compare the pivotal mechanism with the Loehman–Whinston mechanism in the laboratory, where human subjects played robots programmed to reveal their preferences. Preferences of subjects were not controlled by using the induced value method. They used two types of public goods, which were of no clear value to the subjects. Furthermore, the subjects were deceived about the situation. Therefore, it is not possible to draw conclusions about the performance of the pivotal mechanism based on this experiment.

Tideman (1983) reports field experiments in college fraternities, using the pivotal mechanism for real collective decisions. First, as in field experiments it is impossible to control the subjects' preferences. Second, dominant strategies were explained to the subjects, in which case we do not know whether the mechanism itself can induce the subjects to reveal their true preferences without prompting. Third, some of the initial endowments went to the fraternity, which redistributed the money afterwards. This clearly distorted the incentives of the mechanism. In the questionnaire, 21% of the subjects reported overstating their preferences, while 46% reported understating their preferences. Without induced value, it is difficult to evaluate the performance of the pivotal mechanism, such as the extent and magnitude of misrevelation.

Attiyeh, Franciosi and Isaac (forthcoming) report the first well-controlled laboratory experiments on the pivotal mechanism. They reported results from eight independent sessions under two different designs. Design I consisted of a group of five subjects. Design II consisted of a group of ten subjects by replicating the economy of Design I. In each of ten periods the subjects participated in a public goods provision decision-making task. It cost zero to produce the public good, which was of a fixed size. The collective decision was binary, to produce the good or not. The pivotal mechanism was compared to that of a majority rule. Individual values for the public good were sometimes negative and sometimes positive, redrawn each period from the same uniform distribution. Four striking results came from Attiyeh, Franciosi and Isaac (forthcoming):

1. Misrevelation: very few subjects reveal their true valuations. About 10% of the separate bids in Design I and 8% of the separate bids in Design II were truthfully revealing their values.
2. Pattern of misrevelation:
   (a) Positive values: overbid on low values and underbid on high values;
   (b) Negative values: underbid on low values and overbid on high values;
   (c) Bids are closest to value for medium high and medium low draws.
3. Efficiency: 70% of the decisions were efficient. This did not exceed the efficiency of majority rule (also 70%).
4. No-Learning: there was no convergence tendency towards value revelation. This result is similar to the experimental results on second-price auctions, where learning to bid one's true valuation shows little evidence of occurring with experience (e.g., Kagel, 1995, p. 511).

These results raise many questions that should lead to further study of the VCG mechanisms in the public goods context. The failure of most subjects to reveal their true values suggests that the dominant strategy is not transparent.

A recent study by Kawagoe and Mori (2001) analyzes the weakness of incentive compatibility of the pivotal mechanism as a cause for misrevelation. As in Attiyeh, Franciosi, and Isaac (2000), they study the pivotal mechanism in the context of a binary decision-making task to determine whether or not to produce a public project of a fixed size. They conjecture that misrevelation might be due to the fact that the pivotal mechanism is only weakly dominant strategy incentive-compatible. That is, within certain ranges of the strategy space an agent can be indifferent between truth-telling and other strategies. Therefore, an agent might not be able to find the unique dominant strategy without comprehensive understanding of the entire payoff structure. They suggest that one could overcome the problem of weak incentive compatibility by giving the subjects more information about the payoff structure.

Their design has three information treatments. In the Non-Enforcement treatment, each subject was assigned a fixed value and the mechanism was explained without a payoff table. In the Wide Enforcement treatment, each subject was randomly assigned values each round and the mechanism was explained without a payoff table, which is very similar to Attiyeh, Franciosi, and Isaac (2000). In the Deep Enforcement treatment, each subject was assigned a fixed value and given a detailed payoff table. The percentage of truthfully revealing bids was 17% in the Non-Enforcement treatment, 14% in the Wide Enforcement treatment and 47% in the Deep Enforcement treatment. The percentage of public project realized (i.e., efficient decisions) was 40% in the Non-Enforcement treatment, 70% in the Wide Enforcement treatment and 90% in the Deep Enforcement treatment. Overall, more detailed information about the payoff structure significantly improved the rate of dominant strategy play.

Kawagoe and Mori (2001) identified one aspect of the pivotal mechanism, which might have lead to misrevelation. Apart from the weakness of incentive compatibility, the pivotal mechanism provides very little incentives for the subjects to learn their dominant strategies over time. The incentives to learn are provided by connecting non-equilibrium behavior with the resulting losses. In the binary version of the pivotal mechanism, an agent is rarely pivotal in a relatively large economy. Therefore, even if an agent submitted non-equilibrium strategies, her payoff is hardly affected. Note that in the non-binary version of the VCG mechanisms, i.e., when the public goods level is continuous, an agent's message is much more likely to affect the total level of production. Therefore, a non-equilibrium message will result in tax that affects an agent's payoff. Furthermore, strictly convex preferences and continuous levels of public goods are necessary and sufficient for strict incentive compatibility (Kawagoe and Mori, 1998). It would be very interesting to see whether the continuous VCG mechanism has better performance in the laboratory.

3. Nash-efficient Mechanisms

The Groves–Ledyard mechanism Groves and Ledyard (1977) is the first mechanism in a general equilibrium setting, whose Nash equilibrium is Pareto optimal. The mechanism balances the budget both on and off the equilibrium path, but it does not implement Lindahl allocations. Later on, more game forms have been discovered, which implement Lindahl allocations in Nash equilibrium. These include Hurwicz (1979), Walker (1981), Tian (1989), Kim (1993), Peleg (1996), and Chen (2002). Falkinger (1996) presents a mechanism whose Nash equilibrium is Pareto optimal when a parameter is chosen appropriately, however, it does not implement Lindahl allocations and the existence of equilibrium can be delicate.

Most of the experimental studies of Nash-efficient mechanisms focus on the Groves–Ledyard mechanism. Chen and Tang (1998) also compare the Walker mechanism with the Groves–Ledyard mechanism. Falkinger et al. (2000) study the Falkinger mechanism. In all studies except Harstad and Marrese (1982) and Falkinger et al. (2000) quasilinear preferences were used to get a unique Nash equilibrium.[3]

Smith (1979a) reports the first sets of experiments studying various public goods mechanisms. He compared the performance of a voluntary contribution mechanism, a simplified version of the Groves–Ledyard mechanism,[4] and the Smith Auction. The process used in the simplified Groves–Ledyard mechanism was the Smith process, where all the subjects have the opportunity to simultaneously reconsider their messages and to repeat the same choices three times in a row to finalize the production of public goods, and they were paid when agreement was reached. The simplified Groves–Ledyard mechanism provided significantly more public goods than the voluntary contribution mechanism. In the five-subject treatment (R1), one out of three

[3] Bergstrom, Simon, and Titus (1983) show that the Groves–Ledyard mechanism can have a large number of Nash equilibria messages in a general environment. Each will yield an efficient allocation. So far there has been no experimental or theoretical work addressing the equilibrium selection problem in the Groves–Ledyard mechanism.

[4] This simplified version only balanced the budget in equilibrium.

sessions converged to the stage game Nash equilibrium. In the eight-subject replication with different parameters (R2), neither session converged to the Nash equilibrium prediction.

Harstad and Marrese (1981) compare the simplified version of the Groves–Ledyard mechanism under two different processes: the Smith process and the Seriatim process. The Seriatim process also requires unanimity of the subjects to produce the public good, but it differs from the Smith process in that subjects reconsider messages sequentially and only need to repeat their messages once for an iteration to end. They found that only three out of twelve sessions attained approximately Nash equilibrium outcomes.

Harstad and Marrese (1982) study the complete version of the Groves–Ledyard mechanism in Cobb–Douglas economies with the Seriatim process. In the three-subject treatment, one out of five sessions converged to the Nash equilibrium. In the four-subject treatment, one out of four sessions converged to one of the Nash equilibria. This is the only experiment which studied the Groves–Ledyard mechanism in an environment with multiple equilibria, but the equilibrium selection problem was not addressed.

Mori (1989) compares the performance of a Lindahl process with the Groves–Ledyard mechanism.[5] He used a dynamic process similar to the Smith process except that the process stops when each subject repeats her messages once. He ran five sessions for each mechanism, with five subjects in each session. The aggregate levels of public goods provided in each of the Groves–Ledyard sessions were much closer to the Pareto optimal level than those provided using a Lindahl process. On the individual level, each of the five sessions stopped within ten rounds when every subject repeated the same messages. However, since individual messages must be in multiples of .25 while the equilibrium messages were not on the grid, convergence to Nash equilibrium messages was approximate.

None of the above experiments study the effects of the punishment parameter,[6] γ, on the performance of the mechanism. It turns out that this punishment parameter plays an important role in the convergence and stability of the mechanism.

Chen and Plott (1996) first assessed the performance of the Groves–Ledyard mechanism under different punishment parameters. They found that by varying the punishment parameter the dynamics and stability changed dramatically. For a large enough γ, the system converged to its stage game Nash equilibrium very quickly and remained stable; while under a small γ, the system did not converge to its stage game Nash equilibrium. This finding was replicated by Chen and Tang (1998) with more independent sessions (twenty-one sessions: seven for each mechanism) and a longer time series (100 rounds) in an experiment designed to study the learning dynamics. Chen and Tang (1998) also studied the Walker mechanism in the same economic environment.

[5] I thank Toru Mori for providing the data and payoff tables for his experiments, Toshiji Kawagoe and Yo Nagai for the English translation.

[6] Roughly speaking, the punishment parameter in the Groves–Ledyard mechanism determines the magnitude of punishment if a player's contribution deviates from the mean of other players' contributions.

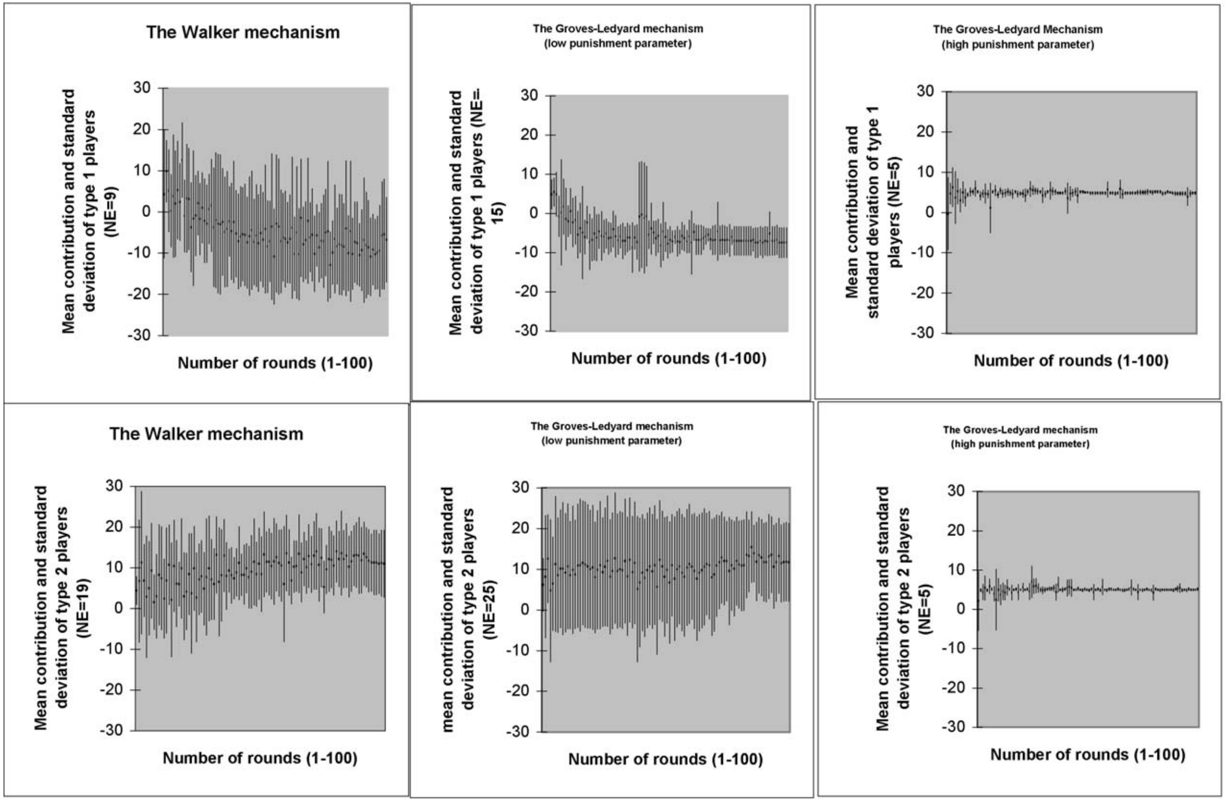

Figure 1. Each column presents the time series data under each mechanism from Chen and Tang (1998). Neither the Walker mechanism nor the Groves–Ledyard mechanism under a low punishment parameter converged to their stage game Nash equilibrium. The Groves–Ledyard mechanism under a high punishment parameter converged very quickly to its stage game Nash equilibrium and remained stable.

Figure 1 presents the time series data from Chen and Tang (1998) for two out of five types of players.[7] Each graph presents the mean (the black dots) and standard deviation (the error bars) for each of the two different types averaged over seven independent sessions for each mechanism – the Walker mechanism, the Groves–Ledyard mechanism under a low punishment parameter (hereafter GL1), and the Groves–Ledyard mechanism under a high punishment parameter (hereafter GL100). From these graphs, it is apparent that the Groves–Ledyard mechanism under a high punishment parameter converged very quickly to its stage game Nash equilibrium and remained stable, while the same mechanism did not converge under a low punishment parameter; the Walker mechanism did not converge to its stage game Nash equilibrium either.

Because of its good dynamic properties, GL100 had far better performance than GL1 and Walker, evaluated in terms of system efficiency, close to Pareto optimal level of public goods provision, less violations of individual rationality constraints and convergence to its stage game equilibrium. All these results are statistically highly significant.

These results illustrate the importance to design mechanisms, which not only have desirable static properties but also good dynamic stability properties. Only when the dynamics lead to the convergence to the static equilibrium, can all the nice static properties be realized.

Muench and Walker (1983) provide a sufficient condition for the Groves–Ledyard mechanism to converge under the Cournot best response dynamics with parameterized quasi-linear utility functions. Chen (1997) provides a necessary and sufficient condition for the GL mechanism to be a supermodular game given quasilinear preferences. Supermodular games are games in which each player's marginal utility of increasing her strategy rises with increases in her rival's strategies, so that (roughly) the player's strategies are "strategic complements." Supermodular games have very robust stability properties, in the sense that a large class of interesting learning dynamics converges to the set bounded by the largest and the smallest Nash equilibrium strategy profiles. This includes Bayesian learning, fictitious play, adaptive learning, Cournot best response and many others (Milgrom and Roberts, 1990). Therefore, Chen (1997) generalizes Muench and Walker's (1983) results to a more general class of preferences and a much wider class of learning dynamics.

OBSERVATION 1 (Chen, 1997). The Groves–Ledyard mechanism is a supermodular game for any $e \in E^Q$ if and only if $\gamma \in [-\min_{i \in N}\{\frac{\partial^2 v_i}{\partial y^2}\}n, +\infty)$.

Therefore, when the punishment parameter is above the threshold, a large class of interesting learning dynamics converge, which is consistent with the experimental results.

Falkinger et al. (2000) study the Falkinger mechanism in a quasilinear as well as a quadratic environment. In the quasilinear environment, the mean contributions moved

[7] The data for the remaining three types of players are not displayed due to limited space, but are available from the author upon request. They display very similar patterns.

towards the Nash equilibrium level but did not quite reach the equilibrium. In the quadratic environment the mean contribution level hovered around the Nash equilibrium, even though none of the 23 sessions had a mean contribution level exactly equal to the Nash equilibrium level in the last five rounds. Therefore, Nash equilibrium was a good description of the average contribution pattern, although individual players did not necessarily play the equilibrium. It is interesting to note that in the quadratic environment the game is *very close* to being a supermodular game: the threshold subsidy coefficient for the mechanism to be supermodular is one, while in the experiment it was set to $2/3$. This is the only experimental study of Nash efficient public goods mechanisms that I am aware of, where parameters were set so close to the threshold.

The following observation organizes all experimental results on Nash-efficient mechanisms with available parameters, by looking at whether they are supermodular games.

OBSERVATION 2 (Chen, 1997). (1) None of the following experiments is a supermodular game: the Groves–Ledyard mechanism studied in Smith's (1979a) R2 treatment, Harstad and Marrese (1982), Mori (1989), Chen and Plott's (1996) low γ treatment, and Chen and Tang's (1998) low γ treatment, the Walker mechanism in Chen and Tang (1998) and Falkinger et al. (2000).

(2) The Groves–Ledyard mechanism under the high γ in Chen and Plott (1996) and Chen and Tang (1998) are both supermodular games.

These past experiments serendipitously studied supermodular mechanisms. The parameters were set either far away from the supermodular threshold (e.g., Chen and Plott, 1996; Chen and Tang, 1998) or very close to the threshold (e.g., Falkinger et al., 2000). None of the experiments systematically varies the parameters from below, close to, at and above the threshold to assess the effects of supermodularity on learning dynamics.

Two recent papers systematically study supermodular mechanisms. Arifovic and Ledyard (2003) conduct computer simulations of an individual learning model in the context of a class of the Groves–Ledyard mechanisms. They vary the punishment parameter systematically, from extremely small to extremely high. They find that their model converges to Nash equilibrium for all values of γ. However, the speed of convergence does depend on the value of the parameter. The speed of convergence is U-shaped: very low and very high values of γ require long periods for convergence, while a range of intermediate values requires the minimum time. In fact, the optimal punishment parameter identified in the simulation is much lower than the supermodularity threshold proposed in Observation 1. Since these results rely on the particular learning model used for the simulation, a natural next step would be to test the prediction in the laboratory with human subjects.

Chen and Gazzale (2002) is the first systematic experimental study of supermodular mechanisms, in the context of the compensation mechanisms. Results of this study is reviewed in Section 4.2.

Four specific game forms[8] implementing Lindahl allocations in Nash equilibrium have been introduced, Hurwicz (1979), Walker (1981), Kim (1993) and Chen (2002). All four improve on the Groves–Ledyard mechanism in the sense that they all satisfy the individual rationality constraint in equilibrium. While Hurwicz (1979) and Walker (1981) can be shown to be unstable for any decentralized adjustment process in certain quadratic environments,[9] the Kim mechanism is stable under a gradient adjustment process given quasilinear utility functions, which is a continuous time version of the Cournot–Nash tatônnement adjustment process. Whether the Kim mechanism is stable under other decentralized learning processes is still an open question.

OBSERVATION 3 (Chen, 1997). None of the Hurwicz (1979), Walker (1981) and Kim (1993) mechanisms is a supermodular game for any $e \in E^Q$. Chen (2002) provides a family of mechanisms, which are supermodular games for $e \in E^Q$.

Since supermodularity is sufficient but not necessary for convergence, this implies that:

1. supermodular mechanisms ought to converge to Nash equilibrium predictions fairly robustly, such as the Groves–Ledyard mechanism under a high punishment parameter;
2. mechanisms which are not supermodular could also converge to its equilibrium under some learning algorithms.

The open question is what learning rules are reasonable and descriptively accurate. Supermodularity provides a sufficient but not necessary condition for convergence under a wide range of learning dynamics. For a complete characterization of the stability of incentive compatible mechanisms, we need a sufficient and necessary condition for convergence under a wide range of learning dynamics, which remains to be found.

4. Mechanisms Using Refinements of Nash as Implementation Concepts

Mechanisms in this category use refinements of Nash equilibrium as the prediction of outcomes. Since there are multiple Nash equilibria, the convergence question is more complex. The behavioral question is whether individuals adopt strategies which support particular refinements of Nash equilibria.

4.1. Perfect Nash Mechanisms

Bagnoli and Lipman (1989) propose a very natural and simple mechanism, which fully implements the core in undominated perfect equilibria in an environment with one private good and a single unit of public good. In a complete information economy agents

[8] Since Tian (1989) and Peleg (1996) do not have specific mechanisms, we will only investigate the supermodularity of these four mechanisms.

[9] See Kim (1986).

voluntarily contribute any non-negative amount of the private good they choose and the social decision is to provide the public good if and only if contributions are sufficient to pay for it. The contributions are refunded otherwise. This result is extended to a public good with finitely many values, where a sequential game with several rounds of contributions implements the core in successively undominated perfect equilibria.

The mechanism in the single unit case is also called the provision-point mechanism in the experimental literature. (See Davis and Holt, 1993 and Ledyard, 1995 for a description of the mechanism and review of related public goods experiments.)

Bagnoli and McKee (1991) test the theoretical predictions of the one unit case in the laboratory. They reported results from seven five-person groups and two ten-person groups. All sessions were implemented as repeated games of 14 periods. Three different treatments were implemented among the five-subject groups: (1) a baseline treatment with homogeneous endowments and homogeneous valuations of the public good (one group), (2) homogeneous endowments and heterogeneous valuations (three groups), and (3) homogeneous endowments and homogeneous valuations (three groups). The two ten-person groups explored only treatments (1) and (2). The results of the Bagnoli and McKee (1991) study provide strong support for the hypothesis that groups will voluntarily contribute sufficient resources to provide the public good, and that group contributions will exactly equal the provision point. Pooling all five-person groups, the public good was provided in 86.7% of the rounds; the Pareto efficient outcome of contributions equal the provision point was observed 54.1% of the cases.

Note, however, there were very few independent sessions in each treatment of the Bagnoli and McKee (1991) study, which raised the question of the robustness of their results. Mysker, Olson, and Williams (1996) report a set of experiments designed to check the robustness of the Bagnoli and McKee (1991) results. They note that Bagnoli and McKee (1991) ran several independent groups simultaneously in the same room and publicly posted contributions for all groups. Bagnoli and McKee (1991) did not use a personal record sheet on which subjects logged their personal contributions and payoffs in each round. Mysker, Olson, and Williams (1996) used a design identical to the Bagnoli and McKee (1991) treatment (1) with homogeneous endowments and homogeneous public good valuation, which also addressed the behavioral effects of the two procedural modification – multiple, simultaneous groups vs single, isolated groups, and the use of a personal record sheet. The Mysker, Olson, and Williams (1996) data challenged the robustness of the Bagnoli and McKee (1991) results. In the Bagnoli and McKee (1991) study the efficient equilibrium contribution is a modal distribution, while in the Mysker, Olson, and Williams (1996) study contributions are evenly distributed along the strategy space.

Both studies have relatively few independent observations (at most three groups) for each treatment. We cannot conclude from the small body of data whether the provision-point mechanism studied in Bagnoli and Lipman works in a laboratory.

Bagnoli, Ben-David, and McKee (1992) report an experiment designed to test the provision-point mechanism for the multiple unit case. The design repeated the single unit case in Bagnoli and McKee (1991) over a sequence. They conducted two treatments: (1) subjects were reassigned to different groups between periods for 6–8 periods, and (2) subjects remained in the same group for the entire session of 15 periods. They found limited support for the prediction that subjects will play equilibrium strategies that achieve a core allocation. This result challenges the strong refinement criterion used in the theory paper. It could also result from the inability of the subjects to coordinate among the multiple equilibria.

Apart from the laboratory experiments, there have also been field experiments using variants of the provision-point mechanism, e.g., the Green Choice Program in upstate New York (Schulze, 1995). In field experiments the complete information assumption in Bagnoli and Lipman (1989) does not hold anymore. Properties of the mechanism under incomplete information are unclear. Other details of the mechanism have been studied as well, for example the effect of refund rules (Isaac, Schmidtz, and Walker, 1989), of rebate rules (Marks and Croson, 1998), of subject pool (Cadsby and Maynes, 1998a, 1998b), of identifiability of the contributors (Croson and Marks, 1998), of incomplete information about valuations (Marks and Croson, 1999), of incomplete information about the number of players or cost of the public good (Rondeau, Schulze, and Poe, 1999), of recommending contributions (Croson and Marks, 1999a), of valuations for the public good (Croson and Marks, 1999b), and of sequentiality of contribution (Coats and Gronberg, 1996). The simplicity of the provision-point mechanism justifies future research into the properties of the mechanism.

4.2. Subgame Perfect Mechanisms

Varian (1994a, 1994b) introduces a class of simple two-stage mechanisms, known as the compensation mechanisms, which implement efficient allocations as subgame-perfect equilibria for economic environments involving externalities and public goods. The basic idea is that each player offers to compensate the other for the "costs" incurred by making the efficient choice.

Andreoni and Varian (1999) report a series of experiments on a variation of the compensation mechanisms. In a particularly simple and elegant design, they considered a modified Prisoners' Dilemma game in which each agent can offer to pay the other agent to cooperate. The mechanism was implemented as card games. Each session consisted of two phases. The first phase was a baseline Prisoners' Dilemma game, called the "Push–Pull" game, which was run for 15 rounds with each subject playing against a different subject each round. The second phase added a contracting stage to the Prisoners' Dilemma, called the "Pay for Push" game, which was run for 25 rounds. They conducted a total of six independent sessions, with eight players per session.

The data show that the mechanism is largely successful at inducing cooperation. The amount of cooperation doubled (from 25.8% to 50.5%) during the second phase when the mechanism was implemented. Players made offers of side payments that should induce cooperation about 63.5% of the time. When such offers were received, subjects responded with cooperation nearly 70% of the time. Full efficient equilibrium was achieved in about 60% of the time. Interestingly, they also found subjects' tastes for cooperation and equity interact significantly with the incentives of the mechanism.[10]

Cheng (1998) studies the dynamic stability of the compensation mechanisms. He proved that the original mechanism is globally stable under the Cournot best response dynamics, but is not supermodular. He also proposed a generalized version of the compensation mechanism, which is supermodular. This generalized version is an excellent playground for studying supermodular mechanisms, as it has two free parameters, one for each type of players. One parameter "determines" whether the mechanism is supermodular, while the other does not play a role in this distinction. This gives the experimenter more freedom in choosing "varying degrees of supermodularity."

Chen and Gazzale (2002) experimentally study the generalized version of the compensation mechanism. They systematically vary the free parameter from below, close to, at and beyond the threshold of supermodularity to assess the effects of supermodularity on the performance of the mechanism. They have three main findings. First, in terms of proportion of equilibrium play and efficiency, they find that supermodular and "near supermodular" mechanisms perform significantly better than those far below the threshold. This finding is consistent with previous experimental findings. Second, they find that from a little below the threshold to the threshold, the improvement in performance is statistically insignificant. This result is important, as theory is silent on this issue. This implies that the performance of "near supermodular" mechanisms, such as the Falkinger mechanism, ought to be comparable to supermodular mechanisms. Therefore, the mechanism designer need not be overly concerned with setting parameters that are firmly above the supermodular threshold – close is just as good. This enlarges the set of robustly stable mechanisms. The third finding concerns the selection of mechanisms within the class of supermodular mechanisms. Again, theory is silent on this issue. Chen and Gazzale find that within the class of supermodular mechanisms, increasing the parameter far beyond the threshold does not significantly improve the performance of the mechanism. Furthermore, increasing another free parameter, which is not related to whether or not the mechanism is supermodular, does improve convergence. Simulation shows that these experimental results persist in the long run.

5. Other Mechanisms

The "Smith Auction" (1979a, 1979b, 1980) was a dynamic process designed for public goods provision. In this mechanism each agent first submit a bid and a proposed quantity of the public good. Each agent's tentative share of the unit cost is the unit cost minus the sum of other agents' bid. The tentative quantity of public good is the average of all agents' proposed quantities. Then each agent is given the right to veto or agree to his/her tentative share of the unit cost and the tentative quantity of the public goods.

[10] For a related theoretical analysis of altruism on the efficiency of public good mechanisms, see Clark (1999).

Group agreement prevails if and only if all players unanimously agree on accepting all the offers. The theoretical properties of this mechanism are not completely understood. However, leaving aside the equilibria of the supergame, the Smith Auction implements the Lindahl allocations in perfect Nash equilibrium in a "one-shot" static game (Banks, Plott, and Porter, 1988).

Smith (1979a) reports 12 sessions of auction experiments in three different environments, where valuation functions were quasilinear, with parameters such that zero contribution would be a dominant strategy for some but not all subjects in a voluntary contribution mechanism. In most sessions average proposals and bid sums were posted on every trial. Nine out of twelve sessions converged to the Pareto optimal quantity of public goods. However, only a small fraction of subjects (11 out of 67) converged to their Lindahl equilibrium bids.

Smith (1980) examines the Smith Auction in environments with income effects. Out of a total of 29 sessions only two sessions failed to reach agreement to provide units of the public good. All sessions that reached agreement provided much larger quantities of public goods than the free-rider quantity. The quantities were not predicted exactly by the Lindahl equilibrium quantity. In fact, the mean quantity is slightly larger than the Lindahl quantity. Compared to the results of Smith (1979a), with income effects, "the LE quantity of the public good is a fair predictor, and the LE bids a very poor predictor of experimental outcomes." (Smith, 1980, p. 597)

Banks, Plott, and Porter (1988) study the Smith Auction and a voluntary contribution mechanism both with and without an additional unanimity feature. They used quasilinear utility functions, where it was a strictly dominant strategy for each agent to contribute zero in a voluntary contribution mechanism. This experiment confirmed that the Smith Auction generated higher levels of public goods provision than the voluntary contribution mechanisms. Aggregate quantity of public goods was near the Pareto optimum. Interestingly, they also found that the inclusion of unanimity reduced the overall efficiency of the process, as well as the success rate of the mechanism. A puzzling result is that overall provision level (efficiency) fell with repetition for the Smith Auction without unanimity, while repetition did not significantly affect the efficiency of the Smith Auction with unanimity. This "no-learning" or decay result could be due to the fact that each session only lasts between three and nine periods.

6. Concluding Remarks

From the existing experimental studies of incentive-compatible mechanisms we can draw several conclusions:

1. Mechanisms with the same static properties can have dramatically different dynamic stability performances.
2. Under the pivotal mechanism it is possible to provide environments where misrevelation is prevalent. More information about the payoff structure help reduce the degree of misrevelation.

3. Available data show that supermodular mechanisms, such as the Groves–Ledyard mechanism under a high punishment parameter, converge robustly well to the Nash equilibrium. Mechanism close to the threshold of being supermodular converges reasonably well. Existing experiments on non-supermodular Nash mechanisms suggest that they did not converge well to the Nash equilibrium predictions.
4. Experiments and simulation show that the performance of "near supermodular" mechanisms ought to be comparable to supermodular mechanisms. Furthermore, within the class of supermodular mechanisms, increasing the parameter far beyond the threshold does not significantly improve the performance of the mechanism.
5. Performance of mechanisms using refinements of Nash as solution concepts is ambiguous. Experimental results could be sensitive to procedural specifications.

Existing experiments on incentive-compatible mechanisms have focused on a few mechanisms. The data, combined with theoretical investigation, give us a fresh perspective on implementation among real people. For these mechanisms that have been studied in the laboratory, we need more replications and robustness tests. Meanwhile, many mechanisms have never been tested in the laboratory. Equilibrium selection problem is almost untouched in this context. Therefore, more experiments should be conducted in this exciting field.

Acknowledgements

I thank Charles Plott for extensive comments on an earlier version, Mark Bagnoli, Rachel Croson, Robert Gazzale, John Kagel, John Ledyard, Bart Lipman, John Morgan, Vernon Smith, James Walker, Arlington Williams for helpful comments.

References

Andreoni, J., Varian, H. (1999). "Pre-play contracting in the prisoners' dilemma". Proceedings of the National Academy of Science 96, 10933–10938.

Arifovic, J. Ledyard, J. (2003). "Computer testbeds and mechanism design: Application to the class of Groves–Ledyard mechanisms for provision of public goods". Working paper, Computing in Economics and Finance 244.

Attiyeh, G., Franciosi, R., Isaac, M. (2000). "Experiments with the pivot process for providing public goods". Public Choice 102, 93–112.

Bagnoli, M., Lipman, B. (1989). "Provision of public goods: Fully implementing the core through private contributions". Review of Economic Studies 56, 583–602.

Bagnoli, M., McKee, M. (1991). "Voluntary contribution games: Efficient provision of public goods". Economic Inquiry 29, 351–366.

Bagnoli, M., Ben-David, S., McKee, M. (1992). "Voluntary provision of public goods: The multiple unit case". Journal of Public Economics 47, 85–106.

Banks, J., Plott, C., Porter, D. (1988). "An experimental analysis of unanimity in public goods provision mechanisms". Review of Economic Studies 55 (2), 301–322.

Bergstrom, Theodore, Simon, Carl P., Titus, Charles J. (1983). "Counting Groves–Ledyard equilibria via degree theory". Journal of Mathematical Economics 12 (2), 167–184.

Cadsby, C., Maynes, E. (1998a). "Choosing between a socially efficient and free-riding equilibrium: Nurses versus economics and business students". Journal of Economic Behavior and Organization 37, 183–192.
Cadsby, C., Maynes, E. (1998b). "Gender and free riding in a threshold public goods game: Experimental evidence". Journal of Economic Behavior and Organization 34, 603–620.
Chen, Y. (1997). "Supermodularity of Nash-efficient public goods mechanisms". University of Bonn SFB303 working paper. (Later circulated under the title, "Dynamic stability of Nash-efficient public goods mechanisms: Reconciling theory and experiments.").
Chen, Y. (2002). "A family of supermodular Nash mechanisms implementing Lindahl allocations". Economic Theory 19, 773–790.
Chen, Y., Gazzale, R. (2002). "Supermodularity and convergence: An experimental study of the compensation mechanism". Manuscript, University of Michigan.
Chen, Y., Plott, C.R. (1996). "The Groves–Ledyard mechanism: An experimental study of institutional design". Journal of Public Economics 59, 335–364.
Chen, Y., Tang, F.-F. (1998). "Learning and incentive compatible mechanisms for public goods provision: An experimental study". Journal of Political Economy 106 (3), 633–662.
Cheng, J. (1998). "Essays on designing economic mechanisms". Ph.D. thesis, University of Michigan.
Clark, J. (1999). "The effects of altruism on the efficiency of public good mechanisms". Working paper, University of Canterbury.
Clarke, Edward H. (1971). "Multipart pricing of public goods". Public Choice 11 (1), 17–33.
Coats, J., Gronberg, T. (1996). "Provision of discrete public goods: An experimental investigation of alternative institutions". Working paper, Texas A&M University.
Corchon, Luis, Wilkie, Simon (1996). "Double implementation of the ratio correspondence by a market mechanism". Review of Economic Design 2 (1), 325–337.
Croson, R., Marks, M. (1998). "Identifiability of individual contributions in a threshold public goods experiment". Journal of Mathematical Psychology 42, 167–190.
Croson, R., Marks, M. (1999a). "Equilibrium selection: Preplay communication and learning". Working paper, OPIM, The Wharton School, University of Pennsylvania.
Croson, R., Marks, M. (1999b). "Step returns in threshold public goods: A meta- and experimental analysis". Working paper, OPIM, The Wharton School, University of Pennsylvania.
Davis, D., Holt, C. (1993). "Experimental Economics". Princeton University Press, Princeton, NJ.
Falkinger, J. (1996). "Efficient private provision of public goods by rewarding deviations from average". Journal of Public Economics 62, 413–422.
Falkinger, J., Fehr, E., Gächter, S., Winter-Ebmer, R. (2000). "A simple mechanism for the efficient provision of public goods – Experimental evidence". American Economic Review 90, 247–264.
Green, J., Laffont, J.-J. (1977). "Characterization of satisfactory mechanisms for the revelation of the preferences for public goods". Econometrica 45, 427–438.
Groves, Theodore (1973). "Incentives in teams". Econometrica 41 (40), 617–631.
Groves, T., Ledyard, J. (1977). "Optimal allocation of public goods: A solution to the 'free rider' problem". Econometrica 45 (4), 783–809.
Groves, T., Ledyard, J. (1987). "Incentive compatibility since 1972". In: Groves, T., Radner, R., Reiter, S. (Eds.), Information, Incentives and Economic Mechanisms: Essays in Honor of Leonid Hurwicz. University of Minnesota Press, Minneapolis.
Groves, Theodore, Loeb, Martin (1975). "Incentives and public inputs". Journal of Public Economics 4 (3), 211–226.
Harstad, R.M., Marrese, M. (1981). "Implementation of mechanism by processes: Public good allocation experiments". Journal of Economic Behavior and Organization 2, 129–151.
Harstad, R.M., Marrese, M. (1982). "Behavioral explanations of efficient public good allocations". Journal of Public Economics 19, 367–383.
Hurwicz, L. (1972). "On informationally decentralized systems". In: McGuire, C., Radner, R. (Eds.), Decision and Organization. North-Holland, Amsterdam, pp. 297–336.
Hurwicz, Leonid (1975). "On the existence of allocation systems whose manipulative Nash equilibria are Pareto optimal". Presented at the Third World Congress of the Econometric Society, Toronto.

Hurwicz, L. (1979). “Outcome functions yielding Walrasian and Lindahl allocations at Nash equilibrium points”. Review of Economic Studies, 217–225.

Isaac, R. Mark, Schmidtz, David, Walker, James M. (1989). “The assurance problem in a laboratory market”. Public Choice 62 (3), 217–236.

Jackson, Matthew O., Moulin, Herve (1992). “Implementing a public project and distributing its cost”. Journal of Economic Theory 57, 125–140.

Kagel, J. (1995). “Auctions: A survey of experimental research”. In: Kagel, J., Roth, A. (Eds.), Handbook of Experimental Economics. Princeton University Press.

Kaneko, Mamoru (1977). “The ratio equilibrium and a voting game in a public goods economy”. Journal of Economic Theory 16 (2), 123–136.

Kawagoe, T., Mori, T. (2001). “Can the pivotal mechanism induce truth-telling? An experimental study”. Public Choice 108, 331–354.

Kawagoe, T., Mori, T. (1998). “A short report on pivotal mechanism experiment”. Mimeo, Nagoya City University.

Kim, T. (1986). “On the nonexistence of a stable Nash mechanism implementing Lindahl allocations”. Mimeo, University of Minnesota.

Kim, T. (1993). “A stable Nash mechanism implementing Lindahl allocations for quasi-linear environments”. Journal of Mathematical Economics 22, 359–371.

Ledyard, J. (1995). “Public goods: A survey of experimental research”. In: Kagel, J., Roth, A. (Eds.), Handbook of Experimental Economics. Princeton University Press.

Ledyard, John O., Palfrey, Thomas R. (1994). “Voting and lottery drafts as efficient public goods mechanisms”. Review of Economic Studies 61, 327–355.

Mailath, G., Postlewaite, A. (1990). “Asymmetric information bargaining problems with many agents”. Review of Economic Studies 57, 351–367.

Mark, L., Matthews, S. (2000). “Dynamic voluntary contribution to a public project”. Review of Economic Studies 67, 327–358.

Marks, M., Croson, R. (1998). “Alternative rebate rules in the provision of a threshold public good: An experimental investigation”. Journal of Public Economics 67, 195–220.

Marks, Melanie B., Croson, Rachel T.A. (1999). “The effect of incomplete information and heterogeneity in the provision point mechanism of voluntary contributions: An experimental investigation”. Public Choice 99 (1–2), 103–118.

Milgrom, P., Roberts, J. (1990). “Rationalizability, learning and equilibrium in games with strategic complementarities”. Econometrica 58 (6), 1255–1277.

Mori, T. (1989). “Effectiveness of mechanisms for public goods provision: An experimental study”. Economic Studies 40 (3), 234–246.

Muench, T., Walker, M. (1983). “Are Groves–Ledyard equilibria attainable?” Review of Economic Studies 50, 393–396.

Mysker, M., Olson, P., Williams, A. (1996). “The voluntary provision of a threshold public good: Further experimental results”. Research in Experimental Economics 6, 149–163.

Peleg, B. (1996). “Double implementation of the Lindahl equilibrium by a continuous mechanism”. Economic Design 2, 311–324.

Roberts, J. (1979). “Incentives and planning procedures for the provision of public goods”. Review of Economic Studies 46, 283–292.

Rondeau, Daniel, Schulze, William D., Poe, Gregory L. (1999). “Voluntary revelation of the demand for public goods using a provision point mechanism”. Journal of Public Economics 72 (3), 455–470.

Scherr, B., Babb, E. (1975). “Pricing public goods: An experiment with two proposed pricing systems”. Public Choice 21, 35–53.

Schulze, W. (1995). “Green pricing: Solutions for the potential free-rider problem”. Mimeo, University of Colorado.

Smith, V. (1979a). “Incentive compatible experimental processes for the provision of public goods”. In: Smith, V. (Ed.), Research in Experimental Economics, vol. 1. JAI Press Inc., Greenwich, CT.

Smith, V. (1979b). "An experimental comparison of three public goods decision mechanisms". Scandinavian Journal of Economics 81, 198–251.

Smith, V. (1980). "Experiments with a decentralized mechanism for public goods decision". American Economic Review 70, 584–599.

Tian, G. (1989). "Implementation of the Lindahl correspondence by a single-valued, feasible, and continuous mechanism". Review of Economic Studies 56, 613–621.

Tideman, T.N. (1983). "An experiments in the demand revealing process". Public Choice 41, 387–402.

Varian, Hal R. (1994a). "Sequential contributions to public goods". Journal of Public Economics 53 (2), 165–186.

Varian, Hal (1994b). "A solution to the problems of externalities when agents are well-informed". American Economic Review 84 (5), 1278–1293.

Vickrey, William S. (1961). "Counterspeculation, auctions, and competitive sealed tenders". Journal of Finance 16, 8–37.

Walker, M. (1980). "On the impossibility of a dominant strategy mechanism to optimally decide public questions". Econometrica 48, 1521–1540.

Walker, M. (1981). "A simple incentive compatible scheme for attaining Lindahl allocations". Econometrica 49, 65–71.