

PARTIAL EQUILIBRIUM APPROACH TO THE FREE-RIDER PROBLEM

Jerry GREEN, Elon KOHLBERG, Jean-Jacques LAFFONT*

*Department of Economics, Harvard University, Cambridge, MA 02138, U.S.A.
Ecole Polytechnique, 75005 Paris, France*

Received September 1975, revised version received May 1976

Groves and others have shown that truthful answers concerning preferences for public goods can be elicited as dominant strategies if appropriate tax-subsidies rules are applied. This paper studies the statistical properties of the total revenues generated by one of the mechanisms with this desirable dominant strategy property. For particular specifications, it is shown that the expected total revenues increase like \sqrt{N} . Further, if these revenues are rebated equally to the individuals in the model, then, although the dominant strategy property no longer obtains, the optimal responses converge to the truth in large economies. As a result, it can be shown that the mechanism with rebates will make the Pareto optimal decision with arbitrarily high probability, provided that the willingness to pay distribution is assumed to be symmetric.

1. Introduction

Even though the problem of manipulating preferences occurs in private goods economies [see Hurwicz (1972), Guesnerie and Laffont (1975)], it is fair to say on the basis of Roberts and Postlewaite (1976), that in situations with a large number of participants, the problem is progressively less severe. On the contrary, with public goods, the failure of decentralized allocation mechanisms requires the creation of decision mechanisms which face the possibility of manipulation with increasing acuteness as the number of agents increases [see Hurwicz (1975), Ledyard and Roberts (1974)]. Since Wicksell's pioneering work (1896), this 'free-rider' problem has been emphasized by many authors [Samuelson (1954), Musgrave (1959)] and remains a fundamental issue of public economics. Indeed, Gibbard (1973) and Satterthwaite (1975) have shown that it is impossible to design Pareto optimal nondictatorial social decision rules when no restrictions are placed on individuals' announced preferences.

However, as in the case of Arrow's impossibility theorem, one may obtain less pessimistic outcomes in more specialized contexts. Further, these results assume that individuals have perfect information about the workings of these

*The authors are grateful to D. McFadden, G. Chichilnisky, R. Guesnerie and the referees for helpful comments. This research was supported by National Science Foundation grants SOC71-03803 to Harvard University and SOC74-11446 to Stanford University.

mechanisms and the actions of others, as well as infinite computational ability. New results may stem from the imposition of meaningful and reasonable constraints on agents' ability to predict the consequences their actions will have on the course of the public decision-making process.

In the case of the determination of the level of provision of public goods and monetary transfers among individuals, Groves (1973, 1974), Groves and Loeb (1975), assuming transferable utility and preferences that are separable in income and the public goods, found a class of mechanisms with the properties that stating one's true preferences is a dominant strategy for each individual and that the outcome is Pareto optimal. In Green and Laffont (1975b) we have shown that these mechanisms are the only ones with these characteristics. A principle drawback of these models is that they do not automatically balance the government budget. The sum of monetary transfers over the set of individuals is zero only in rare, fortuitous, cases. Presumably these deficits or surpluses are balanced by taxation schemes outside this decision-making mechanism. These transfers are not foreseen by the participants when they formulate their original actions.

In view of these limitations, Groves and Ledyard (1975) have constructed a family of mechanisms which, in presence of public goods, yield Pareto optimal outcomes in a truly general equilibrium framework without the assumption of transferable utility. To reach this goal they must, by virtue of the characterization theorem of Green-Laffont (1975b), weaken the result of Groves (1973) in terms of manipulation. The truthful revelation of marginal willingness to pay for public goods occurs at a Nash equilibrium; in other words they must give up the attractive dominant-strategy property which removes all the problems of a noncooperative game theoretic nature. Also, the necessity of a tâtonnement procedure to reach an equilibrium suggests that the deepest form of manipulation is concerned with manipulation along the dynamic process leading to the equilibrium. This idea is pursued in Hurwicz (1975). In any case, the results obtained by Groves and Ledyard (1975) should be considered as a crucial breakthrough since they provide the first general equilibrium treatment of the free-rider problem. However, we believe that it is also important to extend our knowledge of the original Groves mechanisms (1973) which offer considerable potentialities of applications because of the dominant strategy property [see Green and Laffont (1974, 1975a)]. In this paper we present results concerned with the extent to which the application of these mechanisms might cause problems of a general equilibrium nature.

More specifically, we focus attention on one specific mechanism of the Groves class which has the following desirable features. Although the sum of transfer payments among individuals is not exactly zero, if we conceive of individual willingnesses-to-pay as being successive independent samples from a fixed population, it can be shown that the expected per capita net transfer payments go to zero in large economies. The precise result is that the expected transfers grow only at the rate of the square root of the population size.

The implication of this result is that the failure of this mechanism to be self-contained and truly general equilibrium in nature will not be a serious flaw in a large numbers context. If we take the approach of Green and Laffont (1974) and take a random sample of the population, use the mechanism on them, and make a decision for the entire group on the basis of this outcome, the expected net transfer payments received by the unsampled individuals becomes small if the sampled group grows at the same rate as the population size or more slowly. If a small sample is taken, this asymptotic feasibility result is even stronger. On the other hand, one could think of modifying the mechanism by distributing any net transfers equally across the entire population, after having elicited the response of all individuals. This modification would, according to Green and Laffont (1975a), destroy the dominant strategy property since the individuals would foresee the potential influence of their professed willingness-to-pay on their share of the aggregate net transfers, as well as on the decision making process of the original mechanism. However, if we endow individuals with the same beliefs concerning the distribution of tastes in the population, it can be shown that the optimal statement for any individual converges to the truth in large numbers situations. Thus, the bias induced in the decision-making process by redistributing the results of the mechanism in this way is of vanishing importance. On either interpretation, therefore, this paper shows that the desirable dominant strategy property can be essentially preserved in large economies by choosing a particular mechanism of the Groves class. Finally we show that this property of asymptotic dominance of the truthful response implies that the mechanism will give rise to the Pareto optimal decision with arbitrarily high probability as the number of agents increases.

2. The mechanism

We consider an economy with N agents in which a decision is to be made concerning a single well-defined public project. It is assumed, for simplicity, that the project is costless to undertake. The preferences of a typical individual, i , are describable by a utility function which depends on the public decision concerning the project and any monetary transfers, positive or negative, that he might receive, in the following simple way,

$$\begin{aligned} u_i &= v_i + t_i, & \text{if the project is accepted,} \\ u_i &= t_i, & \text{if the project is rejected,} \end{aligned}$$

where t_i is the transfer received by individual i and v_i is a constant which we refer to as his willingness-to-pay.

The willingnesses-to-pay, though known by the individuals, are unknown to the planner, or government. The planner would like to undertake the project if

and only if $\sum v_i \geq 0$. This is equivalent to maximizing a social welfare function with equal weights for all individuals.

The goal of this line of inquiry is to explore the possibility of establishing rules for making transfers among individuals, contingent on their professed willingnesses-to-pay, in such a way as to make telling the truth about these parameters in the private interest of each individual.

More formally, we can let the decision space be $\{0, 1\}$ – ‘0’ for rejection and ‘1’ for acceptance. The outcome space is then $\{0, 1\} \times \mathbf{R}^N$, where (d, t) is the typical element. The i th component of t , t_i , is the transfer received by the i th individual.

A *mechanism* is a mapping, f , from \mathbf{R}^N to $\{0, 1\} \times \mathbf{R}^N$. The domain, \mathbf{R}^N , is the product of the strategy spaces of the individuals, that is, an individual can profess to have a willingness-to-pay equal to any real number. Let $w = (w_1, \dots, w_N)$ be a typical element of the domain. Let $f_0(w)$ be the decision taken when w is the joint strategy, that is, it is the projection of $f(w)$ into $\{0, 1\}$; and let $f_i(w)$ be the transfer that the mechanism assigns to the i th individual.

In a straightforward way, we can write the utility function as it depends on the mechanism and the strategies of all individuals as

$$u_i(f(w)) = f_i(w) + v_i, \quad \text{if } f_0(w) = 1,$$

$$u_i(f(w)) = f_i(w), \quad \text{if } f_0(w) = 0.$$

If $\xi_{-i} \in \mathbf{R}^{N-1}$ and $\xi_i \in \mathbf{R}$, let (ξ_{-i}, ξ_i) represent the vector in \mathbf{R}^N obtained by interjecting ξ_i in the i th component. That is,

$$\xi_{-i} = (\xi_1, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_N),$$

and

$$(\xi_{-i}, \xi_i) = (\xi_1, \dots, \xi_{i-1}, \xi_i, \xi_{i+1}, \dots, \xi_N).$$

We shall say that a mechanism is *strongly individually incentive compatible* (s.i.i.c.) if

$$u_i(f(w_{-i}, v_i)) \geq u_i(f(w_{-i}, w_i))$$

for all $w_{-i} \in \mathbf{R}^{N-1}$ and $w_i \in \mathbf{R}$. In such mechanisms, the truth, v_i , is a *dominant strategy*, for the i th individual.

A mechanism is said to be *successful* if it is s.i.i.c. and if

$$f_0(w) = 0 \quad \text{iff} \quad \sum v_i < 0.$$

In Green and Laffont (1975a) it is shown that the only successful mechanisms

are of the form proposed by Groves (1973) and Groves and Loeb (1973). That is,

$$f_0(w) = 0 \quad \text{iff} \quad \sum w_i < 0,$$

and

$$f_i(w) = (\sum w_{-i} + h_i(w_{-i})), \quad \text{if} \quad \sum w_i \geq 0,$$

$$f_i(w) = h_i(w_{-i}), \quad \text{if} \quad \sum w_i < 0,$$

for some set of arbitrary functions $h_i: \mathbf{R}^{N-1} \rightarrow \mathbf{R}, i = 1, \dots, N$, where $\sum w_{-i} = \sum_{j \neq i} w_j$.

These mechanisms are highly attractive since they are immune to any non-cooperative manipulation, and they achieve optimal allocations of resources but they still suffer from several drawbacks. One is the possibility of manipulation by groups of individuals – a difficult potential problem which we will not treat herein. A second defect is that the sum of the transfer payments is not zero. One can quite easily show that no matter what set of functions $\{h_i\}, i = 1, \dots, N$, is chosen, the resulting sum of transfer payments will not be constant over alternative possible values of $v_i, i = 1, \dots, N$. This problem, which in related literature [see Hurwicz (1972) and references therein] is called ‘feasibility’, is certainly a serious one. It means that the mechanism does not really define an allocation of resources at all since it does not specify how any surplus collected should be redistributed nor how any deficit should be financed.

For example, if we concentrate on the particular mechanism defined by $h_i \equiv 0, i = 1, \dots, N$, then, when the project is accepted, the sum of all transfers received is $(N-1)\sum w_i$. That is, the total transfer payment is $(N-1)$ times the social value of the project itself. On practical grounds, this is quite clearly too much of a distortion. One might think of subtracting a suitably large constant function h , but this would result in heavy taxes in the case of rejection and would be similarly unacceptable.

We use, therefore, the following mechanism, which will be referred to as the *pivotal mechanism* for reasons that will be made apparent below. The function h is defined by

$$h_i(w_{-i}) = -\sum w_{-i}, \quad \text{if} \quad \sum w_{-i} \geq 0,$$

$$h_i(w_{-i}) = 0, \quad \text{if} \quad \sum w_{-i} < 0.$$

The transfer function is accordingly:

$$f_i(w) = 0, \quad \text{if} \quad \sum w_i \geq 0 \text{ and } \sum w_{-i} \geq 0,$$

$$f_i(w) = \sum w_{-i}, \quad \text{if} \quad \sum w_i \geq 0 \text{ and } \sum w_{-i} < 0,$$

$$f_i(w) = -\sum w_{-i}, \quad \text{if} \quad \sum w_i < 0 \text{ and } \sum w_{-i} \geq 0,$$

$$f_i(w) = 0, \quad \text{if} \quad \sum w_i < 0 \text{ and } \sum w_{-i} < 0.$$

Several features of this mechanism should be mentioned. First, it taxes individuals but never subsidizes them. The individuals who pay the tax are those whose stated willingnesses-to-pay are such as to change the sign of the aggregate – these individuals are pivotal in the decision, or *pivots*. Their payment is just the amount by which the decision that their vote causes is damaging to the remainder of the agents. This mechanism is the public goods analogue of the bidding process first proposed by Vickrey (1961). Second, the nature of these payments is such that no individual would prefer not to participate in the mechanism, allowing the others to make the decision instead. This follows because the maximum tax that could be enforced upon him is at most the amount that changing the social decision would be worth to him. Moreover it is clear that any attempt to raise additional revenues through the mechanism by making h_i lower at any value of w_{-i} would run the risk of causing some agents to refuse to participate in the process. If they could not be compelled to do so, as might be the case if a sampling approach were taken, this would bias the results of the mechanism and would render it unsuccessful. Third, and finally, the aggregate tax collection is bounded at a much lower level than for other mechanisms. It is easy to show that the upper bound on the sum of payments by pivotal individuals is one half the sum of the absolute value of the stated willingnesses-to-pay. This can be attained only when $\sum w_i$ is exactly zero, so that the project is accepted according to our sign convention, and all individuals whose willingness-to-pay is positive have stated identical values of that parameter. This bound is not always lower than the value of $(N-1)\sum w_i$ as can be seen, in particular, in the case in which the bound is achieved as mentioned above. It is therefore incorrect to claim that the pivotal mechanism is ‘cheaper’ or ‘closer-to-being feasible’ than $h = 0$, or any other particular mechanism. However, it can be shown that the pivotal mechanism generates, in a statistical sense, a much lower value of total payments.

3. Expected payments by pivotal individuals

In light of the comments made above, we assume that the actual willingnesses-to-pay in the population of N individuals can be viewed as having arisen from repeated independent sampling from a fixed distribution. The spirit of this assumption is that this mechanism will be used to evaluate many potential public projects whose characteristics are similar, and the decision-making agency, or government, is interested in minimizing the extent of the necessary transfer payments in the long run. We therefore consider the expected level of payments made by the pivotal individuals, since only these individuals are taxed in the mechanism under consideration. By symmetry, it suffices to consider a single agent and compute his expected tax payments.

Let $F(\cdot)$ be the distribution of the willingness-to-pay of a single agent, v . Let $F_{N-1}(\cdot)$ be the distribution of the sum of $N-1$ such independent random

variables, which we denote by x . The agent will be pivotal under the mechanism described in section 2, if $|x| < |v|$ and either

$$v < 0 < x$$

or

$$x < 0 < v.$$

Whenever he is pivotal his payment is $|x|$.

Therefore the expected total taxes collected from all the agents are

$$E_N = N \int_{-\infty}^{+\infty} dF(v) \int_{\{|x| < |v| \text{ and } xv < 0\}} |x| dF_{N-1}(x). \tag{1}$$

In the case in which $F(\cdot)$ is a unimodal distribution with mean and mode at zero and finite variance σ^2 , it can be shown¹ that E_N/\sqrt{N} converges to $\sigma/2\sqrt{(2\pi)}$, and in particular that the per capita expected payment is going to zero. The convergence of this expression in the general mean zero case is an open question. Below, assuming that the distribution $F(\cdot)$ is normal with mean zero, we give an explicit calculation of E_N and its limiting properties.

$$\begin{aligned} E_N &= 2N \int_{-\infty}^0 dF(v) \int_0^{-v} x dF_{N-1}(x) \\ &= \frac{N}{\pi\sqrt{(N-1)}} \int_{-\infty}^0 \exp[-v^2/2] dv \int_0^{-v} x \exp[-x^2/2(N-1)] dx \\ &= \frac{-N\sqrt{(N-1)}}{\pi} \int_{-\infty}^0 \exp[-v^2/2] dv [\exp[-v^2/2(N-1)] - 1] \\ &= \frac{1}{\sqrt{(2\pi)}} [N\sqrt{(N-1)} - (N-1)\sqrt{N}]. \end{aligned}$$

It follows that

$$\begin{aligned} \sqrt{(2\pi)} \frac{E_N}{N} &= \sqrt{(N-1)} - (N-1) \frac{\sqrt{N}}{N} \\ &= \frac{1}{\sqrt{N}} (\sqrt{((N-1)N)} - (N-1)) \\ &\leq \frac{1}{\sqrt{N}} (N - (N-1)) = \frac{1}{\sqrt{N}} \end{aligned}$$

¹See Green and Laffont (1976). In that paper, a proof that no Groves mechanism can have in general a balanced budget is also given.

If F had a nonzero mean, μ , then E_N/N would converge to zero even more rapidly than $1/\sqrt{N}$. We demonstrate this in the next section via a simulation when F is normal with mean $\mu \neq 0$. The convergence is exponential. Intuitively the reason is that x is centered around $(N-1)\mu$ which is very large when N is large. It is therefore extremely unlikely that v would dominate x in absolute value and have the opposite sign. A proof of this result under the assumption that $N^2 f_{N-1}(x)$ converges uniformly to zero as N goes to infinity, where $f_{N-1}(\cdot)$ is the density of F_{N-1} , is given in Green and Laffont (1976).

4. Simulations

The expressions obtained in section 3 were evaluated using well-known methods of numerical analysis. In fig. 1 we show the values of $E_N(\mu)$ for $|\mu| = 0, 0.1, 0.2, 0.3$. It is clear that $E_N(\cdot)$ converges pointwise as $\mu \rightarrow 0$, and that each $E_N(\mu)$ achieves a maximum beyond which it decreases rapidly as $N \rightarrow \infty$. For each $E_N(0)$, however, the computations confirm the fact that it increases on the order of $N^{1/2}$ as $N \rightarrow \infty$. Results for selected larger values of N and $|\mu|$ are gathered in table 1.

As a practical matter, the mean zero cases, in which our results are weakest, are of central importance. When the planner is certain about the population mean and believes it to be nonzero, the social decision in a large economy is almost surely determined in advance of the application of any mechanism. We therefore concentrate on the zero mean case throughout our further analysis in section 5.

5. Asymptotic dominance

The interpretation of the asymptotic feasibility results given above follows the spirit of the approximation results found in a variety of contexts in Walrasian general equilibrium theory. That is, if an imbalance becomes negligible on a per capita basis in a large economy, then the system is approximately in equilibrium. All measurements, both physical and economic, being imprecise, we can therefore regard the system as being in a true, exact, equilibrium.

In the current model, that is the case. The taxes collected are negligible on a per capita basis and the dominant strategy property insures that truthful responses are always being given. However, it is not clear that we should be completely satisfied with such an approximation in this context. Unlike the case of individuals demanding private goods at prices which depart slightly from the true equilibrium, in the present model the incentive to behave optimally grows progressively weaker with the size of the economy. This raises the following query. Suppose that the taxes collected were rebated equally to all individuals. Would individuals still have the incentive to tell the truth?

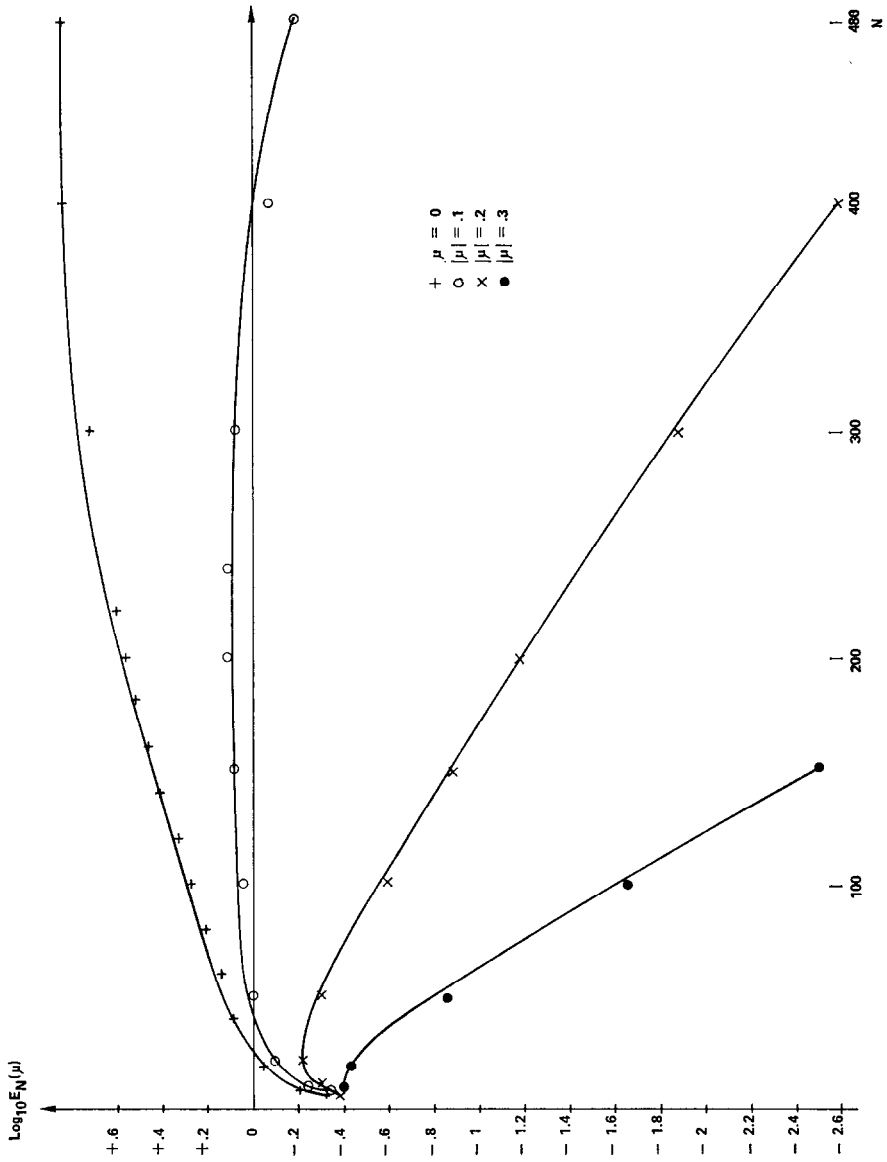


Fig. 1. Normal case: $E_N(\mu)$ for various N and μ .

In order to provide a meaningful answer to this question we must specify the expectations each individual holds concerning the distribution of stated willingnesses-to-pay as expressed by the other people. The rebates, which insure the exact feasibility of the process, destroy the dominant strategy property. To keep the analysis parallel to that of the previous sections, we assume that each of the individuals in the society believes that all of the others are drawn independently from a normal population with zero mean. We will then calculate his optimal stated willingness-to-pay and show that this approaches the truth as the economy grows.

It should be pointed out that this is a much stronger result than that obtained in the earlier sections. It is not enough that per capita expected rebates are going to zero. The potential regret for distorting one's answer is also decreasing since it is progressively less likely that each individual is pivotal. One must weigh the expected *change* in the rebate to be received against this regret, and must show that the latter is the stronger effect. As both the rebate and the regret are approaching zero on the order of $N^{-1/2}$, a more refined analysis is necessary before the optimum response can be determined.

Let us first compute the expected utility of any response, w , given that decisions will be taken and taxes will be computed using the pivotal mechanism, and that equal rebates will then be made to all individuals.

For simplicity, we will concentrate on the case in which $w \geq 0$, the case of $w < 0$ being symmetric. Let there be $N+1$ individuals in all; therefore from the viewpoint of any one of them there are N others. The sum of their responses is denoted x , and is assumed to be distributed normally with mean zero and variance N (without loss of generality). Let $F_N(x)$ be the distribution function of this random variable. Then the expected utility derived from the project itself is

$$v \int_{-w}^{\infty} dF_N(x). \quad (2)$$

Since $1/(N+1)$ of one's own taxes are rebated, the expected payments net of this rebate are

$$\frac{N}{N+1} \int_{-w}^0 x dF_N(x), \quad (3)$$

because the individual is pivotal (when $w > 0$) if and only if $-w < x < 0$.

Table 1
Normal case.

μ	N									
	10	20	50	100	200	500	1000	2000	5000	10000
0	0.61	0.89	0.13×10^1	0.18×10^1	0.36×10^1	0.75×10^1	0.87×10^1	0.87×10^1	0.48×10^{-53}	0.82×10^{-216}
0.5	0.18	0.77×10^{-1}	0.27×10^{-2}	0.75×10^{-5}	0.53×10^{-10}	0.57×10^{-26}	0.57×10^{-26}	0.57×10^{-26}	0.25×10^{-107}	0.82×10^{-216}
1	0.53×10^{-2}	0.52×10^{-4}	0.24×10^{-10}	0.48×10^{-21}	0.17×10^{-42}	0.25×10^{-107}	0.25×10^{-107}	0.25×10^{-107}	0.65×10^{-243}	0
1.5	0.14×10^{-4}	0.27×10^{-9}	0.94×10^{-24}	0.52×10^{-48}	0.12×10^{-96}	0.65×10^{-243}	0.65×10^{-243}	0.65×10^{-243}	0	0
2	0.41×10^{-8}	0.10×10^{-16}	0.16×10^{-42}	0.95×10^{-86}	0.22×10^{-172}	0	0	0	0	0

Finally, we have to compute the expected rebates from the N other pivotal payments. The sum of the expected payments by these individuals, if they are pivotal in the affirmative, is

$$N \int_{-\infty}^{-w} dF_{N-1}(x) \int_{-w-x}^{\infty} (w+x) dF(Z), \tag{4}$$

where $F_{N-1}(x)$ is the distribution of the sum of $N-1$ random variables, each of which is distributed normally with zero mean and unit variance – $F(Z)$ being the distribution of the willingness to pay of such a typical potentially pivotal individual. From the individual’s point of view, his statement, w , is known, and only the other N are random. This explains the asymmetry of this expression about zero and the difference between it and those of section 3. Similarly, expected total payments by pivotal individuals who have defeated the public project are

$$-N \int_{-w}^{\infty} dF_{N-1}(x) \int_{-\infty}^{-w-x} (w+x) dF(Z). \tag{5}$$

Therefore, since the individual is entitled to $1/(N+1)$ of these as a subsidy, his expected utility from this source is

$$\begin{aligned} \frac{N}{N+1} & \left[\int_{-w}^{\infty} dF_{N-1}(x) \int_{-\infty}^{-w-x} (w+x) dF(z) \right. \\ & \left. - \int_{-\infty}^{-w} dF_{N-1}(x) \int_{-w-x}^{\infty} (w+x) dF(Z) \right] \tag{6} \end{aligned}$$

The sum of (2) and (6) minus (3) is the expected utility of the response $w > 0$.

It is easy to see that because $1/(N+1)$ of one’s own taxes are rebated, the optimal response considering only (2)–(3) is $[(N+1)/N]v_i$. This is true independently of one’s beliefs about the distribution of responses by the others. However, considering the fact that one’s share in the rebates of others’ pivotal payments is partially determined by one’s own statement, the optimal answer may be somewhat modified.

Consider the case of three individuals with v ’s equal to 1, -4 , and 2 respectively. The first individual is not a pivot. He receives a subsidy of 1 which is his share of the second individual’s pivotal payment. By saying $1\frac{1}{2}$, however, he can increase his subsidy to $1\frac{1}{3}$; similarly, any statement less than 2 will cause the second individual to be pivotal by a larger amount, and hence will increase his payment. However, if he follows this course, but is ignorant of the actual statements to be made by the others, then he may be hurt by this exaggeration. For example, if the third individual would say 3 instead of 2, then by saying $1\frac{1}{2}$ he has made himself pivotal and must pay a tax which would have been avoided by a truthful response. The larger the variability of the sum of everyone else’s answers, the more risky it becomes to try and increase one’s share in total

taxes collected by distorting one's response. For uncertain beliefs about the others' tastes that are symmetric around zero, it can be shown that this consideration induces a response closer to zero than $[(N+1)/N]v_i$.

We will now argue that $w > 0$ is the relevant case to consider when $v > 0$ and by symmetry the optimum will be characterized by $w < 0$ when $v < 0$. The original mechanism, which has a payoff function of $((2)-(3)) \times [(N+1)/N]$ is optimized at $w = v$ by the dominant strategy property. The additional term,

$$-\frac{1}{N+1} \int_{-w}^0 x \, dF_N(x), \tag{7}$$

which enters into (3), represents the expected rebate on one's own pivotal payments. It is clearly a symmetric function around zero which achieves its minimum there and is monotone on either side of this point. Therefore the maximum of (2) and (3) must be at a point v^* which has the same sign as v and is greater in absolute value.

The value of the expression (6) on the other hand is maximized at zero and is a symmetric function which is monotone decreasing on its positive branch and increasing on its negative branch. Therefore the maximal expected utility will be attained between v^* and zero. Therefore, when $v > 0$, the relevant maximand is the expected utility indicated above. Furthermore, the remarks above indicate that the maximum will be attained on the interior of the interval $[0, v^*]$, and will be characterized by a zero of the derivative at that point.

This derivative can be calculated to be

$$\begin{aligned} & v f^N(-w) - \frac{N}{N+1} w f^N(-w) \\ & + \frac{N}{N+1} \left[\int_{-w}^{\infty} dF^{N-1}(x) [\int_{-\infty}^{-w-x} dF(Z) - (w+x)f(-w-x)] \right] \\ & - \frac{N}{N+1} \left[\int_{-\infty}^{-w} dF^{N-1}(x) [(w+x)f(-w-x) + \int_{-w-x}^{\infty} dF(Z)] \right], \tag{8} \end{aligned}$$

that is,

$$\begin{aligned} & v f^N(-w) - \frac{N}{N+1} w f^N(-w) + \frac{N}{N+1} \int_{-\infty}^{\infty} dF^{N-1}(x) \int_{-\infty}^{-w-x} dF(Z) \\ & - \frac{N}{N+1} \int_{-\infty}^{-w} dF^{N-1}(x) - \frac{N}{N+1} \int_{-\infty}^{\infty} dF^{N-1}(x)(w+x)f(-w-x), \tag{9} \end{aligned}$$

where f^N and f are the densities of the distributions F^N and F respectively.

We now assume normally distributed individual willingness-to-pay, with zero mean and unit variance.² In this case, let $[N/(N+1)]g_N(w)$ denote the last term of (9). Then

$$g_N(w) = \frac{1}{\sqrt{(2\pi(N-1))}} \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} \exp[-x^2/2(N-1)](w+x) \times \exp[-(w+x)^2/2] dx,$$

but

$$\frac{x^2}{N-1} + (w+x)^2 = \frac{N}{N-1} \left(x + \frac{N-1}{N} w \right)^2 + \frac{w^2}{N},$$

so that

$$g_N(w) = \frac{1}{\sqrt{(2\pi(N-1))}\sqrt{(2\pi)}} \exp[-w^2/N] \times \int_{-\infty}^{\infty} \exp\left[-\left(x + \frac{N-1}{N} w\right)^2 / N-1/N\right] \left(x + \frac{N-1}{N} w + \frac{w}{N}\right) dx = \frac{1}{\sqrt{(2\pi(N-1))}\sqrt{(2\pi)}} \cdot \exp[-w^2/2N] \cdot \frac{w}{N} \sqrt{\left(2\pi \frac{N-1}{N}\right)}.$$

Clearly,

$$\sqrt{(N)}g_N(w) \text{ converges to zero uniformly in } [0, v^*]. \tag{10}$$

Denote now the sum of the third and fourth terms of (9) by $[N/(N+1)]h_N(w)$. Then

$$h'_N(w) = - \frac{1}{\sqrt{(2\pi(N-1))}\sqrt{(2\pi)}} \int_{-\infty}^{\infty} \exp[-x^2/2(N-1)] \times \exp[-(w+x)^2/2] dx + \frac{1}{\sqrt{(2\pi(N-1))}} \exp[-w^2/2(N-1)] = - \frac{1}{\sqrt{(2\pi(N-1))}\sqrt{(2\pi)}} \int_{-\infty}^{\infty} \exp\left[-\left(x + \frac{N-1}{N} w\right)^2 / 2\left(\frac{N-1}{N}\right)\right] \times \exp[-w^2/2N] dx + \frac{1}{\sqrt{(2\pi(N-1))}} \exp[-w^2/2(N-1)] = - \frac{1}{\sqrt{(2\pi N)}} \exp[-w^2/2N] + \frac{1}{\sqrt{(2\pi(N-1))}} \exp[-w^2/2(N-1)].$$

²This is generalized in Green and Laffont (1976) to unimodal symmetric distributions such that $\sqrt{(N)}f_N(\cdot)$ converges uniformly to a constant as $N \rightarrow \infty$.

It is clear that $\sqrt{(N)}h'_N(w)$ converges uniformly to zero in $[0, v^*]$. Since $h_N(0) = \frac{1}{2} - \frac{1}{2} = 0$, it follows that

$$\sqrt{(N)}h_N(w) \text{ converges uniformly to zero in } [0, v^*]. \tag{11}$$

Let w_N be the maximum point of the expected utility (2)–(3)+(6). Then, setting the derivative (9) equal to zero, we get

$$\left(v - \frac{N}{N+1}w_N\right) \frac{1}{\sqrt{(2\pi N)}} \exp\left[-\frac{w_N^2}{2N}\right] + \frac{N}{N+1} h'_N(w_N) - \frac{N}{N+1} g'_N(w_N) = 0.$$

Multiplying this equation by \sqrt{N} and applying (10) and (11) we get that w_N converges to v , as was to be shown.

6. Asymptotic successfulness

The primary goal of these mechanisms is to induce collective decisions that are Pareto optimal. Closing the system through a rebate of the pivotal payments collected induces a bias in individual responses that is small if the individual perceives himself to be only one among a large number of agents in the system. This does not guarantee that the social decision is ‘approximately optimal.’ Indeed, since we view individuals having stochastically generated tastes, the approximate optimality of such closed mechanisms, if it can be proved to hold at all, must be a statistical statement.

We will say that a mechanism is *asymptotically successful* if for large enough economies it gives rise to the correct decision regarding acceptance or rejection of the project with arbitrarily high probability. If the rebate mechanism analyzed above did not have this property, the fact that individuals can be induced to give asymptotically correct answers would be of little relevance. *A priori* it seems possible that if in order to induce answers of a given degree of accuracy we need a very large number of agents, then the errors made by these agents in responding to the mechanism might accumulate in such a way as to distort the aggregate.

In this section we assume that the v_i are independently and identically distributed with zero mean. Further, it is supposed that the distribution through which they are generated is symmetric about zero and has a density $f(\cdot)$.

In light of the preceding section, we know that the optimal response, $w_N(\cdot)$, of an agent believing that the population is of size N satisfies;

- (a) $w_N(\cdot)$ is an integrable function;
- (b) $w_N(v)$ approaches v as N goes to infinity, for every v ;

(c) $|w_N(v) - v| < |v|$;

(d) $w_N(v) = -w_N(-v)$.

To prove that the mechanism is asymptotically successful it suffices to show that

$$\Pr \left[\sum_i^N v_i \geq 0 \text{ and } \sum_i^N w_N(v_i) \leq 0 \right] \quad (12)$$

goes to zero as N approaches ∞ .

As a result, the probability of taking a non-optimal decision approaches zero with increasing N .

Theorem. Under the hypotheses above, the probability that the mechanism will take the correct decision approaches one as the number of agents goes to infinity.

Proof. Note first that (12) is equivalent to

$$\Pr \left[\frac{\sum_{i=1}^N v_i}{\sqrt{N}} \geq 0 \text{ and } \frac{\sum_{i=1}^N w_N(v_i)}{\sqrt{N}} \leq 0 \right] \rightarrow 0 \text{ when } N \rightarrow \infty. \quad (13)$$

Denote:

$$\hat{v}_{iN} = \frac{v_i}{\sqrt{N}},$$

$$\hat{v}_N = \sum_{i=1}^N u_{iN},$$

$$\hat{w}_{iN} = \frac{w_N(v_i)}{\sqrt{N}},$$

$$\hat{w}_N = \sum_{i=1}^N \varepsilon_{iN},$$

$$y_N^* = \begin{bmatrix} \hat{v}_N \\ \hat{w}_N \end{bmatrix}.$$

Let $y = [y^1, y^2]$ be the vector of jointly normally distributed random variables with zero mean and variance-covariance matrix

$$\begin{bmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 \end{bmatrix}.$$

According to the theorem of Varadarajan [see Rao (1965)] it suffices to show that for all $(\lambda_1, \lambda_2) \in R \times R$,

$$\lambda_1 \hat{v}_N + \lambda_2 \hat{w}_N \rightarrow \lambda_1 y^1 + \lambda_2 y^2,$$

in order to conclude that

$$y_N \rightarrow y.$$

That is, $\lambda_1 \hat{v}_N + \lambda_2 \hat{w}_N$ approaches a univariate normal with zero mean and variance $(\lambda_1 + \lambda_2)^2 \sigma^2$.

Let

$$a_{iN} = \lambda_1 \hat{v}_{iN} + \lambda_2 \hat{w}_{iN},$$

and let F_{iN} be the distribution of a_{iN} . Finally, let

$$a_N = \lambda_1 \hat{v}_N + \lambda_2 \hat{w}_N.$$

We will apply the Lindeberg-Feller central limit theorem [see Malinvaud (1969)] to conclude that a_N converges to a normal distribution for every choice of λ_1 and λ_2 , and that the variance of this distribution is the same as that of the random variable $\lambda_1 y^1 + \lambda_2 y^2$, namely, $(\lambda_1 + \lambda_2)^2 \sigma^2$.

Let

$$\sigma_\infty^2 = \lim_N \text{var } a_N.$$

The Lindeberg-Feller theorem says that if

$$\sigma_\infty^2 > 0, \tag{14}$$

and, for all $\bar{a} > 0$,

$$\sum_{i=1}^N \int_{|a| > \bar{a}} a^2 dF_{iN} \rightarrow 0, \text{ as } N \rightarrow \infty, \tag{15}$$

then a_N approaches a normal random variable with mean zero and variance σ_∞^2 .

First we verify (14), proving in particular that $\sigma_\infty^2 = (\lambda_1 + \lambda_2)^2 \sigma^2 > 0$. Then we verify (15). By the remarks above this will complete the proof of the theorem.

Let $w_N(v_i) - v_i = e_N(v_i)$. Note that, by (d) and symmetry of the distribution of v_i , $Ee_N(v_i) = 0$. Thus

$$a_{iN} = \frac{1}{\sqrt{N}} [(\lambda_1 + \lambda_2)v_i + \lambda_2 e_N(v_i)],$$

$$\sigma_{iN}^2 = \frac{1}{N} [(\lambda_1 + \lambda_2)\sigma^2 + 2\lambda_2(\lambda_1 + \lambda_2)E v_i e_N(v_i) + \lambda_2^2 E e_N(v_i)^2].$$

Let

$$\begin{aligned}\sigma_N^2 &= \sum_{i=1}^N \sigma_{iN}^2 \\ &= (\lambda_1 + \lambda_2)^2 \sigma^2 + 2\lambda_2(\lambda_1 + \lambda_2) \frac{1}{N} \sum_{i=1}^N E v_i e_N(v_i) + \lambda_2^2 \frac{1}{N} \sum_{i=1}^N E e_N(v_i)^2.\end{aligned}$$

But

$$\frac{1}{N} \sum_{i=1}^N E v_i e_N(v_i) = E v_i e_N(v_i),$$

since the v_i are identically distributed,

$$\int_{-\infty}^{+\infty} v_i e_N(v_i) f(v_i) dv_i = \int_{-\infty}^{+\infty} \Psi_N(v_i) dv_i,$$

with $|\Psi_N(v_i)| < v_i^2 f(v_i)$ by (c) and

$$\Psi_N(v_i) \xrightarrow{N} 0$$

for each v_i by (b).

Hence, by the Lebesgue dominated convergence theorem,

$$E v_i e_N(v_i) \xrightarrow{N} 0.$$

Similarly,

$$\frac{1}{N} \sum_{i=1}^N E e_N(v_i)^2 = E e_N(v_i)^2 \xrightarrow{N} 0;$$

hence,

$$\sigma_N^2 \rightarrow (\lambda_1 + \lambda_2)^2 \sigma^2 = \sigma_\infty^2.$$

To prove that (15) holds, take $\bar{a} > 0$ and consider

$$\int_{|a| > \bar{a}} a^2 dF_{iN} = \int_{A_N} \frac{[\lambda_1 v_i + \lambda_2 w_N(v_i)]^2}{N} dF(v_i),$$

where

$$A_N = \{v_i \mid |(\lambda_1 + \lambda_2)v_i + \lambda_2 e_N(v_i)| > \bar{a}\sqrt{N}\}.$$

Since the v_i are identically distributed, these integrals are independent of i .

The following inequalities show that when multiplied by N they are each still converging to zero. Hence the sum will go to zero as required by (15). Now,

$$|\lambda_1 v_i + \lambda_2 w_N(v_i)|^2 \leq \lambda_1^2 v_i^2 + \lambda_2^2 w_N(v_i)^2 + 2\lambda_1 \lambda_2 |v_i w_N(v_i)| \leq (\lambda_1 + 2\lambda_2)^2 v_i^2$$

by (c).

Hence,

$$\int_{A_N} u^2 dF_{iN} < \frac{(\lambda_1 + 2\lambda_2)^2}{N} \leq \int_{A_N} v_i^2 dF(v_i).$$

Since $\int v_i^2 dF(v_i)$ is finite, it suffices to show that $\bigcap_{N=1}^\infty A_N = \{0\}$ in order that $\int_{A_N} v_i^2 dF(v_i)$ will be made arbitrarily small with N . But $|\lambda_1 v + \lambda_2 w_N(v)| \leq [\lambda_1 + 2\lambda_2]|v|$, and thus for every v , there exists N_0 such that

$$|v| \leq \frac{\bar{a}\sqrt{N}}{\lambda_1 + 2\lambda_2}, \quad \text{for any } N \geq N_0.$$

Thus $\bigcap_{N=1}^\infty A_N = \{0\}$ and for any $\varepsilon > 0$, there exists N such that

$$\int_{A_N} v_i^2 dF(v_i) < \frac{\varepsilon}{(\lambda_1 + 2\lambda_2)^2},$$

or

$$\left| \sum_{i=1}^N \int_{|a| \geq \bar{a}} a^2 dF_{iN} \right| < \varepsilon, \quad \text{for } N > N^*.$$

This verifies (15).

Hence y_N converges in distribution to y and hence, for N sufficiently large,

$$|\Pr[\hat{v}_N \geq 0 \text{ and } \hat{w}_N \leq 0] - \Pr[y^1 \geq 0 \text{ and } y^2 \leq 0]| < \varepsilon.$$

Since

$$\Pr[y^1 \geq 0 \text{ and } y^2 \leq 0] = 0,$$

we have the desired result of asymptotic successfulness of the mechanism.

References

Gibbard, A., 1973, Manipulation of voting schemes: A general result, *Econometrica* 41, 587-601.
 Green, J. and J.-J. Laffont, On the Revelation of Preferences for Public Goods, forthcoming.
 Green, J. and J.-J. Laffont, 1975a, Strength of incentives for individually incentive compatible mechanisms, Harvard Institute of Economic Research, Discussion Paper No. 411.

- Green, J. and J.-J. Laffont, 1975b, Characterization of strongly individually incentive compatible mechanisms for the revelation of preferences for public goods, Harvard Institute of Economic Research, Discussion Paper No. 412, *Econometrica*, forthcoming.
- Green, J. and J.-J. Laffont, 1976, Revelation des preferences pour les biens publics—1ère partie: Carcterisation des mécanismes satisfaisants, Laboratoire d'Econométrie de l'Ecole Polytechnique, A 125-0176, Paris.
- Groves, T., 1973, Incentives in teams, *Econometrica* 41, 617–631.
- Groves, T., 1974, Information, incentives and the internalization of production externalities, Discussion Paper No. 87, Center for Mathematical Studies in Economics and Management Science, Northwestern University.
- Groves, T. and J. Ledyard, 1975, An incentive mechanism for efficient resource allocation in general equilibrium with public goods, Discussion Paper No. 119, Center for Mathematical Studies in Economics and Management Science, Northwestern University.
- Groves, T. and M. Loeb, 1975, Incentives and public inputs, *Journal of Public Economics* 4, no. 3, 211–226.
- Guesnerie, R. and J.-J. Laffont, 1975, Advantageous reallocation of initial resources (CEPREMAP, Paris).
- Hurwicz, L., 1975, Nash equilibrium and Pareto optimality for public and private goods, presented to NSF-NBER Conference on Decentralization, Princeton University.
- Hurwicz, L., 1972, On informationally decentralized systems, in: R. Radner and C.B. McGuire, eds., *Decision and organization* (North-Holland, Amsterdam).
- Ledyard, J. and D.J. Roberts, 1974, On the incentive problem with public goods, Discussion Paper No. 116, Center for Mathematical Studies in Economics and Management Science, Northwestern University.
- Malinvaud, E., 1969, *Statistical methods of econometrics*, second edition (Rand McNally, Chicago).
- Musgrave, R., 1959, *The theory of public finance* (McGraw-Hill, New York).
- Rao, C.R., 1965, *Linear statistical inference and its applications*, second edition (John Wiley, New York).
- Roberts, D.J. and A. Postlewaite, 1973, The incentives for price-taking behavior in large exchange economies, *Econometrica* 44, No. 1.
- Samuelson, P., 1954, The pure theory of public expenditures, *Review of Economics and Statistics* 36, 387–389.
- Satterthwaite, M., 1975, The existence of a strategy proof voting procedure: A topic in social choice theory, Ph.D. dissertation, University of Wisconsin.
- Vickrey, W., 1961, Counterspeculation, auctions and competitive sealed tenders, *Journal of Finance* 41, No. 1.
- Wicksell, K., 1896, *Finanztheoretische Untersuchungen und das Steuerwesen Schwedens*, Jena, Germany, English translation in Musgrave and Peacock, eds., *Classics in the theory of public finance* (I.E.A. Macmillan, London) 1958.