DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

VOTING IS APPROXIMATELY OPTIMAL: A CHARACTERIZATION OF INTERIM EFFICIENCY WITH PUBLIC GOODS

John O. Ledyard

Thomas R. Palfrey
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 956

February 1996 (revised June 1997)

Voting is Approximately Optimal: A Characterization of Interim Efficiency with Public Goods

John O. Ledyard Thomas R. Palfrey

Abstract

We consider a Bayesian public goods environment with independent private valuations, where a public good can be produced at constant returns to scale, up to some capacity. We fully characterize the interim efficient allocation rules and prove that they correspond to decision rules based on a virtual cost-benefit criterion, together with the appropriate incentive taxes. Compared to the classical Lindahl-Samuelson solution there are generally distortions that depend on the welfare weights because the efficient way to reduce the tax burden on low-valuation (resp: high-valuation) types is to reduce (resp: increase) the level of provision of the public good. Second, we explore the implementation of efficient allocations by means of simple, dominant strategy voting rules, called *referenda*. In a referendum, individuals vote for or against production of the public good. If a sufficiently large fraction vote in favor, the good is provided at maximum capacity and costs are distributed equally across the population. Otherwise the good is not produced. We prove that for each interim efficient allocation rule there exists a referendum that approximates that achieves the same total surplus in large populations. Furthermore, if there is common value uncertainty in addition to the private valuations uncertainty, then the approximately optimal referendum is unique.

JEL classification numbers: 024, 026

Key words: public goods, mechanism design, referendum, interim efficiency, incentive compatibility, private values

Voting is Approximately Optimal: A Characterization of Interim Efficiency with Public Goods *

John O. Ledyard Thomas R. Palfrey

1 Introduction

In this paper, we consider the following classical public goods problem. A group of individuals must decide on a level of a public good that is produced according to constant returns to scale up to some capacity constraint. In addition to deciding the level of public good, the group must decide how to tax the individuals in the group in order to cover the cost. The distribution of the burden of taxation is important because different individuals have different marginal rates of substitution between the private good (taxes) and the public good, and may have different incomes as well. These individual marginal rates of substitution are private information; that is, each individual knows his or her own marginal rate of substitution, but not those of the other members of the group. Adopting a Bayesian mechanism design framework, we assume that the distribution of marginal rates of substitution is common knowledge.

We are interested in characterizing efficiency in this environment and are also interested in characterizing those mechanisms that one might expect to actually arise in practice. This suggests two approaches, one from normative considerations and one from positive considerations. On the normative side, we ask: What should an active planner (a mechanism designer) do? A well-known special case of this problem has been solved for one particular social welfare function (e.g. d'Aspremont and Gérard-Varet 1979) that

*We are grateful for the support of the National Science Foundation grant no. SBR-9223701, and of the New Millennium Program of the Jet Propulsion Laboratory of NASA. The second author is grateful for the hospitality and research support at Laboratoire d'Economie Industrielle and at Centre d'Enseignement et de Recherche en Analyse Socio-Economique. The paper has benefited from the comments of seminar participants at Université de la Mediterranée, Northwestern University, Université de Cergy-Pontoise, the Conference on Efficiency in Economics with Public Goods and Private Information at the University of Venice, the Roy Seminar at Ecole des Ponts et Chaussées, the 1996 Francqui Prize Colloquium in Brussels, and from discussions with Louis-André Gérard-Varet and Jean-Charles Rochet. Three referees and a coeditor provided additional helpful suggestions which have improved the paper. The usual disclaimer applies.

is insensitive to the distribution of cost shares. What distinguishes our work here is that we consider a planner who is maximizing a welfare function that is sensitive to the allocation of cost shares over the different valuation-types. Simply put, the planner may care who pays. This is represented formally by type-contingent welfare weights.

Why might the consideration of such distributional goals be relevant? What rationale could ever be given for non-constant welfare weights? Perhaps the simplest example to answer these questions corresponds to public decisions with zero production costs. Such cases are well-approximated in the real world by social legislation such as blue laws, smoking and drinking prohibitions, clothing requirements at beaches, and so forth. Suppose one is considering the implementation of one such social regulation. Many would argue that if implementation takes place, then the losers (i.e. those with a negative valuation to the proposed regulation) should be compensated.[1] But one runs into the (incentive compatibility) problem that if you naively say you are going to compensate all losers, then everyone will claim to be a loser, possibly leading to production never occurring.. A planner might want, therefore, to give some weight to the losers but not to the exclusion of all others. Obviously, in order to compensate the losers in such decisions, incentive taxes need to be carefully constructed which will achieve such type-contingent redistribution, at least to the extent limited by incentive compatibility constraints. As we will show below, there is a direct and intuitive link between the desired degree of such compensation and the corresponding distortions away from the Lindahl-Samuelson optimum. In this particular example, significant compensation of losers would necessitate a corresponding degree of underproduction relative to the classic solution. Other weighting schemes would correspond to other type-distributional goals, and could lead to either under- or overproduction.

A second reason to consider non-constant welfare weights is evident if one concedes that this partial equilibrium model is embedded in a richer general equilibrium structure, where income or wealth distribution is a goal of the planner. If preferences for the public good are correlated with income or wealth in a systematic way, then the public good mechanism can be used as an instrument for redistribution, and unequal welfare weights would be a reflection of the planner's redistributive goals.

A third rationale for unequal weights is more direct. For reasons which may have to do only remotely with issues of compensating losers or wealth redistribution, certain kinds of type-dependent cost-sharing may be deemed desirable on their own merits. A classic example of this is the class of proportional cost-sharing rules, whereby individuals valuing the public good more *should* bear a proportionally larger share of the costs (e.g. Jackson and Moulin 1992). Such normative goals would correspond to a system of welfare weights that decrease in type in a particular way.

For the positive approach to the mechanism design problem, we ask: What would we expect to see in practice? Here we are looking for a concept of efficiency or stability because we would expect inefficient or unstable mechanisms to be replaced by others.

[1] Sometimes, this requirement is implicitly imposed as a voluntary participation constraint.

Under complete information, these concepts correspond to Pareto optimality and the core, respectively. Under asymmetric information the problem is a bit more subtle, and there remains no true consensus on the appropriate equivalent concepts.[2] We therefore take a minimalist approach, and look at a natural extension of Pareto optimality to asymmetric information. In the analysis below we assume that all decisions, including whether to change the mechanism, are made at the interim stage – that is, when each agent knows his or her type, but not anyone else's type. If there is no communication, then the set of interim incentive efficient mechanisms consists of those incentive compatible mechanisms for which it cannot be common knowledge that there is another mechanism which generates a unanimous improvement. We would expect therefore that surviving institutions would be interim incentive efficient.

Luckily we do not have to choose between normative and positive approaches to this problem. As pointed out in Holmström and Myerson (1983), a mechanism is interim efficient if and only if there exist type-dependent social welfare weights for which that mechanism solves the planners optimization problem subject to feasibility and incentive compatibility constraints. Thus, by varying the welfare weights in our planner's problem, we map out the entire set of mechanisms which are interim incentive efficient. Thus, a complete solution to this problem, posed either from a normative or positive standpoint, is equivalent to fully characterizing the set of interim efficient mechanisms for the production of public goods in this framework.

A complete characterization of interim efficiency has been done for the special case where the types are identically distributed and can only take on two values (Ledyard and Palfrey 1994). There it was shown that optimal production always takes a special form in which the public good is provided if and only if the number of high valuation types exceeds a threshold number which depends on the welfare weights and the distribution of types. The greater the welfare weight on high valuation types, the lower the optimal threshold. With more than two types (as in this paper) the optimal mechanism generally depends on the exact profile of types in a more complicated way. In the first half of this paper, we characterize interim efficient mechanisms and obtain some comparative statics about how the optimal mechanism changes with the underlying distribution of types and with the welfare weights of the welfare function.

In the second half of the paper, we look deeper at the positive question of what mechanisms we might expect to see in practice, by investigating the optimality properties of simple, dominant strategy voting mechanisms. We show that in large populations the performance of the optimal mechanisms can be approximated by using binary voting schemes in which the public good is produced (at maximum capacity) if and only if a threshold proportion of "yes" votes is met, and costs are shared equally by all types of all agents. Otherwise there is no production. We call such schemes *referenda.* In particular we show that for every interim efficient mechanism there is a referendum such that the aggregate welfare achieved from the voting scheme converges, as the population grows,

[2]See Holmstrom and Myerson (1983), Cramton and Palfrey (1995), and Crawford (1985) for good discussions of the difficulties of extending these concepts to asymmetric information.

to the aggregate welfare achieved from the interim efficient mechanism. Moreover, if the distribution of valuations is not known precisely by the planner, then the optimal referendum is *uniquely* determined.

2 The Model

There are N people who must decide on the quantity, q of a public good that is produced according to constant returns to scale and has a maximum level $Y = 1$. The cost of producing $q \in [0, 1]$ is equal to Kq. In addition, they must decide how to distribute the production costs. Because of the linear production technology, the optimal level of the public good will always be either 0 or 1, so this is equivalent to a problem of deciding on whether or not to produce a discrete public good. We let a^i denote individuals i's share of the cost, in units of the consumption of the private good, and assume it can take any real value. Therefore the set of feasible levels of production and cost shares are given by

$$(a^1, \ldots, a^N, q) \in \Re^N \times [0, 1]$$

such that

$$\text{with } Kq \leq \sum_{i=1}^{N} a^i.$$

Individual *preferences* are assumed to be risk-neutral and quasilinear in the level of public good production and the taxes (cost shares), so the utility to type v^i of agent i for an allocation (q, a) is given by

$$V^i = v^i q - a^i.$$

Thus, v^i represents the marginal rate of substitution between the public and private good, or "public good valuation" of type v^i. We refer to v^i as player i's "value." We assume that each individual knows his own value, v^i, and does not know the values of the other individuals. We assume that the individual values (v^i) are independently distributed, with the (common knowledge) *cdf* of i's value denoted $F_i(\cdot)$ and the support of F_i is $V^i = [\underline{v}^i, \overline{v}^i]$, where $\underline{v}^i < K/N < \overline{v}^i$. We assume F_i has a continuous positive density on V^i. Note that $\underline{v}^i < 0$ is allowed.

Clearly under these assumptions, our choice of normalization of the utility function is arbitrary up to an affine transformation. In particular, it is equivalent (in terms of individual decision theory) to the models of asymmetric information about contribution costs (a^i), where utilities are normalized[3] so that the marginal utility of the public good

[3]This normalization can be made as long as $\underline{v}^i > 0$.

(v) equals 1. So that $u^i = q - (\frac{1}{v^i})a^i$. However,the class of ex-ante incentive efficient mechanisms (in the sense of Holmström and Myerson 1983) will be different under the two normalizations. [4] So, below, we will focus on the set of interim-incentive efficient mechanisms. That set is independent of whatever (type dependent) normalization one chooses.

A *mechanism* consists of a message space for each agent and an outcome function mapping message profiles into probability distributions over the set of feasible allocations. By the revelation principle, the properties (in terms of allocations) of any optimal mechanism can be duplicated by an incentive compatible, direct mechanism in which the message space for agent i is simply the set of possible types (values) in the support of F_i. A strategy for i is a mapping $\sigma^i : V^i \to V^i$, that is, a decision rule that specifies a reported type for each possible type. We refer to the identity mapping as the truthful strategy. By the linearity of the individual utility functions, there is also no loss in restricting attention to deterministic mechanisms. Thus, we will denote a feasible direct mechanism simply as a function

$$\eta : V^N \longrightarrow \{(a^1, \ldots, a^N, q) \in R^N \times [0,1] | \sum_{i=1}^{N} a^i \geq Kq\}.$$

We denote the public good allocation component of η at type profile v by $q(v)$, and the private good tax for i by $a_i(v)$.

Besides feasibility, the main restriction on η is that it be *incentive compatible*, which means that it is a Bayesian equilibrium of η for all agents to adopt a strategy of truthfully reporting their type. Given a strategy profile $\sigma^i : V^i \longrightarrow V^i$ and a mechanism, η, let the interim utility of type v^i of agent i, assuming all others truthfully report their type, be denoted by:

$$\hat{u}^i(\eta, v^i, \sigma^i) = \int_{V^{-i}} [v^i q(\sigma^i(v^i), v^{-i}) - a^i(\sigma^i(v^i), v^{-i})] dF(v|v^i)$$

Let $u^i(\eta, v^i) = \hat{u}^i(\eta, v^i, I)$ where I denotes the truthful strategy $I(v) = v$. Then η is incentive compatible if and only if $u^i(\eta, v^i) \geq \hat{u}^i(\eta, v^i, \sigma^i)$ for all v^i, σ^i.

The set of interim incentive efficient allocation rules[5] can be represented as the solutions to a set of maximization problems. Let $\lambda > 0$ be a system of welfare weights, a measurable function mapping types into the positive real line, so that $\lambda(v^i)$ represents the welfare weight assigned to type v^i of agent i. Then η is interim efficient if and only

[4]The fact that ex ante efficiency is sensitive to utility normalizations is discussed in Ledyard and Palfrey 1994 (p. 333).

[5]For the remainder, we simply refer to such allocations as "interim efficient."

if there is a λ such that η maximizes $\Sigma_i \int_{\underline{v}^i}^{\overline{v}^i} \lambda_i(v^i)u^i(\eta, v^i)dF_i(v^i)$ over the set of feasible and incentive compatible mechanisms.[6]

We now proceed to characterize this set.

3 The Characterization

As indicated above, we represent interim efficient rules as a solution to a constrained maximization problem. First we need to identify incentive compatible mechanisms in a useful way.

For smooth mechanisms, when preferences are linear, the characterization of incentive compatibility in terms of derivatives is well-known. There are basically two features of such mechanisms. First, an envelope condition is satisfied, namely that the total derivative of the interim utility for i with respect to type when players adopt truthful strategies is equal to the partial derivative with respect to type (i.e., fixing the reports of all agents). Second, the interim utility to i under truthful reporting is convex in i's type. This is stated formally below.

Lemma (Rochet, 1987): *If $\hat{u}^i$ linear in v^i and η is twice continuously differentiable, then η is incentive compatible if and only if*

$$\begin{array}{ll} \nabla_{v^i} u^i(\eta, v^i) = \nabla_{v^i} \hat{u}^i(\eta, v^i, I) & \quad \text{(i)} \\ u^i(\eta, v^i) \textit{ is convex in } v^i. & \quad \text{(ii)} \end{array}$$

For our problem $\nabla_{v^i} u^i(\eta, v^i) = Q^i(v^i) \equiv \int_{V^{-i}} q(v)dF(v \mid v^i)$. So u^i is convex in v^i if and only if $Q^{i\prime}(v^i) \geq 0 \ \forall \ v^i$. Using these facts we can see that a mechanism (q, a) is interim efficient if and only if there is a λ such that (q, a) solves $\max \int_V \sum_i \lambda(v^i)(v^i q(v) - a^i(v))dF(v)$ subject to $0 \leq q(v) \leq 1 \ \forall \ v$, $Q^{i\prime}(v^i) \geq 0 \ \forall \ i, v^i$, $\nabla_{v^i} u^i(v^i) = \nabla_{v^i} \hat{u}^i(v^i, I) \ \forall \ i, v^i$, and $\sum_i a^i(v) = Kq(v) \ \forall \ v$.

Using the approach of Mirrlees (1971) and Wilson (1993) we construct the Lagrangian equivalent problem

$$\max_{\eta} \min_{\psi, \delta} \sum_i \int_{\underline{v}^i}^{\overline{v}^i} \lambda_i(v^i)u^i(\eta, v^i)dF_i(v^i)1$$

$$+\sum_i \int_{\underline{v}^i}^{\overline{v}^i} \psi^i(v^i) \left[u^i_{v^i}(\eta, v^i) - \hat{u}^i_{v^i}(\eta, v^i, I)\right] dv^i$$

$$+\int_V \delta(v) \left[\sum_i a^i(v) - Kq(v)\right] dv$$

[6] See Holmstrom and Myerson (1983)

subject to:

$$\begin{array}{ll} 0 \le q(v) \le 1 & \forall\, v \in V \\ Q^{i\prime}(v^i) \ge 0 & \forall\, i,\; v^i \in V^i \end{array}$$

where ψ^i and γ are multipliers for (first order) incentive compatibility and feasibility, respectively. Applying Green's Theorem and substituting the identity $u^i(\eta, v^i) = \hat{u}^i(\eta, v^i, I)$ converts the maximization problem to:

$$\max_{\eta} \min_{\psi,\delta} \sum_i \int_{\underline{v}^i}^{\overline{v}^i} \Big\{ \hat{u}^i(\eta, v^i, I)[\lambda_i(v^i) f_i(v^i) - \psi_i'(v^i)\Big]$$

$$-\psi_i(v^i)\hat{u}^i_{v^i}(\eta, v^i, I)\Big\}\, dv^i + \int_V \delta(v)\left(\sum_i a^i(v) - Kq(v)\right) dv$$

$$+\sum_i \int_{\partial V^i} \hat{u}^i(\eta, v^i, I) \cdot (\psi_i(v^i)\xi_i(v^i))dv^i$$

subject to:

$$\begin{array}{ll} 0 \le q(v) \le 1 & \forall\, v \in V \\ Q^{i\prime}(v^i) \ge 0 & \forall\, i,\; v^i \in V^i \end{array}$$

where ∂V^i denotes the boundary of V^i and ξ_i points outward at v^i.

We are now in a position to give a complete characterization of the class of interim efficient mechanisms.

Theorem 1 *(q^*, a^*) is an interim efficient mechanism if and only if $\exists\ \lambda \gg 0$ with $\int_{\underline{v}^i}^{\overline{v}^i} \lambda(v^i) dF^i(v^i) = 1\ \forall\ i$, such that*

(a) $\forall v,\ q^*(v)$ maximizes $\{\sum_i w^i(v^i) - K\}$

subject to:

$$\begin{array}{ll} 0 \le q(v) \le 1 & \forall\, v \in V \\ Q^{i\prime}(v^i) \ge 0 & \forall\, i,\; v^i \in V^i \end{array}$$

where

$$w^i(v^i) = v^i - \frac{1 - F_i(v^i)}{f_i(v^i)} + \frac{\int_{v^i}^{\overline{v}^i} \lambda^i(t^i) dF_i(t^i)}{f_i(v^i)}$$

and

$$\text{(b)} \quad a^{*i}(v) = \int_{\underline{v}^i}^{v^i} t^i dQ^{*i}(t^i) + \alpha^i(v)$$

where

$$\sum_i \alpha^i(v) = Kq^*(v) - \sum_i \int_{\underline{v}^i}^{vi} t^i dQ^{*i}(t^i) \;\; \forall v$$
$$\text{and } \frac{\partial}{\partial v^i} \int \alpha^i(v) dF(v \mid v^i) = 0 \;\; \forall \; i, \, v^i.$$

Proof: A sketch is given. For further details see Ledyard and Palfrey (1996).

First notice that the restriction of λ to $\int_{\underline{v}^i}^{\overline{v}^i} \lambda_i(v^i)dF(v^i) = 1 \; \forall \; i$ is without loss of generality. Since utilities are linear in the transfers, for some welfare weights total welfare can be made arbitrarily large simply by making *ex ante* transfers from one individual to another individual. That is, if, for two agents i and j, it were the case that

$$\int_{\underline{v}^i}^{\overline{v}^i} \lambda_i(v^i)dF_i(v^i) < \int_{\underline{v}^j}^{\overline{v}^j} \lambda_j(v^j)dF_j(v^j)$$

then total welfare could be made arbitrarily large by making *ex ante* transfers of the private good from i to j. Thus, a solution to the maximization problem only exists when the welfare weights are, in expectation, the same for all agents. Thus, without loss of generality, we restrict the welfare weights to satisfy

$$\infty > \int_{\underline{v}^i}^{\overline{v}^i} \lambda_i(s)dF_i(s) = 1 \;\; \forall i.$$

We can write (*) as

(**)

$$\max_{q \in T} \min_{\psi, \delta} \int_V \left[\sum_i (\lambda_i(v^i) - \frac{\psi_i'(v^i)}{f_i(v^i)})(v^i q(v) - a^i(v)) - \frac{\psi_i(v^i)}{f_i(v^i)} q(v) \right] dF(v)$$
$$+ \int_V \delta(v) \left(\sum_i a^i(v) - Kq(v) \right) dv + \sum_i \left[\psi^i(\overline{v}^i)\hat{u}^i(\overline{v}^i) - \psi^i(\underline{v}^i)\hat{u}(\underline{v}^i) \right].$$

From the first order conditions with respect to $a^i(v), \delta(v)$, and $\psi^i(v^i)$ we obtain, for $\underline{v}^i < v^i < \overline{v}^i$,

$$(1) \quad -(\lambda_i(v^i)f_i(v^i) - \psi_i'(v^i)) + \gamma(v)f_i(v^i) = 0$$

$$(2) \quad \textstyle\sum_i a^i(v^i) - Kq(v) \begin{cases} = 0 & \text{if } \gamma(v) > 0 \\ \geq 0 & \text{if } \gamma(v) = 0 \end{cases}$$
$$(3) \quad \textstyle\frac{\partial}{\partial v^i} \int_{V^{-i}} a^i(v)dF(v \mid v^i) \equiv A^{i\prime}(v^i) = v^i Q^{i\prime}(v^i) \; \forall \; i, v^i$$

where $\gamma(v) = \delta(v)/f(v)$.

From (1) it follows that γ is constant in v. Integration of (1) gives $\psi^i(v^i) = F_i(v^i)$ $(\lambda_i^-(v^i) - \gamma) + C$ where $\lambda_i^-(v^i)$ is the expected value of λ_i conditional on i's valuation being less than or equal to v^i.

Part (b) of the theorem[7] follows from (2) and (3).

Finally, the continuity of ψ^i along with the first order conditions for a^i at $\underline{v}^i$ and $\overline{v}^i$ imply that $\psi^i(\underline{v}^i) = \psi^i(\overline{v}^i) = 0$. So $C = 0$ and $\gamma \equiv \int_{\underline{v}^i}^{\overline{v}^i} \lambda_i(v^i) dF_i(v^i) = 1$. Substituting all of this into (**) implies that we must find q^* to solve

$$\max_q \int \left[\sum_i (v^i - \frac{\psi_i(v^i)}{f_i(v^i)}) - K \right] q(v) dF(v)$$

$$\textit{subject to}:$$

$$\begin{array}{ll} 0 \leq q(v) \leq 1 & \forall\ v \in V \\ Q^{i\prime}(v^i) \geq 0 & \forall\ i,\ v^i \in V^i \end{array}$$

QED

4 Interpreting the Characterization

Call $w_i(v^i) = v_i - \frac{F_i(v_i)}{f_i(v_i)}(\lambda_i^-(v_i) - 1)$, type v^i of agent i's *virtual valuation* (à la Myerson). Suppose[8] $w_i'(v^i) \geq 0\ \ \forall i, v^i$. Then, since $Q^{i*}(v^i) = prob(\sum_{j \neq i} w_j(v^j) \geq K - w_i(v^i))$ it will be true that $Q^{i\prime}(v^i) \geq 0$ is never binding. So for (λ, F) such that $w_i'(v^i) \geq 0 \forall i, v^i$, interim efficient $q^*(v)$ satisfy

$$\begin{aligned} q^*(v) &= 1 \text{ if } \textstyle\sum_i w^i(v^i) \geq K \\ &= 0 \text{ otherwise.} \end{aligned}$$

This is a *virtual cost-benefit* criterion. The virtual utility has a familiar interpretation (see for example Myerson 1981). It equals the "true" public good valuation of the v^i-type inflated[9] by a factor that depends on the distribution of types and on the welfare weights. The benchmark case is the one where $\lambda_i(v^i) = 1$ for all i and v^i. In this case the first best optimal level of public good is 1 or 0 depending only on whether or not $\sum_i [v^i - \frac{K}{N}] \geq 0$.

[7]The existence of such an α for any given q was first shown by d'Aspremont and Gérard-Varet (1979). One α that satisfies (b1) and (b2) is $\alpha^i(v) = \frac{K}{N}[q^*(v) - Q^{*i}(v^i) + \frac{1}{N-1}\sum_{j \neq i} Q^{*j}(V^j)] - \frac{1}{N-1}\sum_{j \neq i} \int_{\underline{v}^j}^{v^j} s_j dQ^{*j}(s_j)$.

[8]This is the so-called "regular" case, where the second order condition is never binding.

[9]This could be deflated if $\lambda_i^-(v^i) > 1$.

That is, produce if and only if the sum of the marginal rates of substitution exceed the marginal production cost. This is the Lindahl-Samuelson solution, precisely the solution investigated in most previous papers on the optimal provision of public good. (See d'Aspremont and Gérard-Varet 1979). This simplification arises because the allocation of the private good (i.e., the incidence of the costs on different types) does not affect social welfare. For this reason, incentive compatibility does not reduce social welfare relative to the first best solution. However, it must be emphasized that this is a very special case. It is in fact the only system of welfare weights where incentive compatibility does not cause distortions relative to the first best solution.[10]

To better understand the intuition behind the virtual valuations, one can think of the mechanism operating in the following way. Each agent (truthfully) reports a valuation. If the public good is produced, then each agent pays the incentive tax, which equals a constant plus that agent's valuation minus his "informational rent", $\frac{1-F_i(v^i)}{f_i(v^i)}$. Recall from standard incentive theory that this is the amount that can be extracted from an agent, given incentive constraints. Of course, in this public good problem, the objective of the mechanism is not to extract rent from agents, so any excess incentive tax will be distributed lump sum back to the agents, by adjusting the incentive tax by a constant. Thus, if the good is provided, the government spends K to produce the public good and makes a lump-sum refund (which is formally captured by the constant (i.e. independent of v^i) that is added to each agent's incentive tax. The portion of this refund that comes from type v^i of agent i equals $(v^i - \frac{1-F_i(v^i)}{f_i(v^i)} - \frac{K}{N})$. There are two other terms that complete the social cost/benefit picture, as it concerns type v^i of agent i. One is simply that producing the public good, produces a direct benefit of v^i to agent i, which is valued socially as $\lambda_i(v^i)v^i$. Last but not least, is the fact that the incentive tax (before refund) is a social cost, and this social cost equals $\lambda_i(v^i)v^i - \frac{\int_{v^i}^{\overline{v}^i} \lambda_i(t^i)dF_i(t^i)}{f_i(v^i)}$.

Collecting all these terms, gives us type v^i of agent i's contribution to the marginal net social value of producing the public good. Denoting this by $\overline{w}^i(v^i)$, gives us:

$$\begin{aligned}
\overline{w}^i(v^i) &= \lambda_i(v^i)v^i - [\lambda_i(v^i)v^i - \frac{\int_{v^i}^{\overline{v}^i} \lambda_i(t^i)dF_i(t^i)}{f_i(v^i)}] + [v^i - \frac{1-F_i(v^i)}{f_i(v^i)} - \frac{K}{N}] \\
&= v^i - \frac{1-F_i(v^i)}{f_i(v^i)} + \frac{\int_{v^i}^{\overline{v}^i} \lambda_i(t^i)dF_i(t^i)}{f_i(v^i)} - \frac{K}{N} \\
&= w^i(v^i) - \frac{K}{N}
\end{aligned}$$

which is the *cost adjusted* virtual valuation of type v^i of agent i.

[10] Actually, this is the only system of welfare weights in which a first best solution exists. For any other weights, welfare can be arbitrarily increased by shifting the allocation of the private good to one particular type of some individual. Since we impose no feasibility bounds on the allocation of the private good, this means that the first best solution does not exist. Of course, with incentive compatibility constraints, the second-best problem is well defined.

Notice that in the special case of neutral distributional weights, $\int_{\underline{v}^i}^{\overline{v}^i} \lambda_i(t^i)dF_i(t^i) = 1 - F_i(v^i)$, so that $\lambda_i(v^i)v^i - \frac{\int_{v^i}^{\overline{v}^i} \lambda_i(t^i)dF_i(t^i)}{f_i(v^i)} = v^i - \frac{1-F_i(v^i)}{f_i(v^i)}$ and as a result there are no welfare costs associated with charging the incentive taxes in a type-dependent way and then redistributing them back in a lump sum fashion. Otherwise there is a cost to doing this.

The form of virtual utilities also makes it easy to see how distortions away from the classic optimum are related to the welfare weights. For example, if λ_i is decreasing in type then generally the interim efficient solution calls for *underproduction* relative to the Lindahl-Samuelson solution, since $\psi_i(v^i)$ is positive for all types. That is, the virtual valuations are always less than true valuations, so the sum of the true valuations must more than exceed the production cost in order for production to be optimal. Conversely, if λ_i is increasing in type, then there should be *overproduction* relative to the Lindahl-Samuelson solution.

The discussion above assumes monotone virtual utilities, which ensures that maximization of the relaxed program, without the $Q^{i\prime}(v^i) \geq 0$ constraint, automatically satisfies that constraint. It is straightforward to see what is required for virtual utilities to be monotone in type, and this provides a nice intuition for how our results differ from standard incentive problems of this type (e.g. Guesnerie and Laffont 1984). From above,

$$w_i = v^i - \frac{1 - F_i(v^i)}{f_i(v^i)} + \frac{\int_{v^i}^{\overline{v}^i} \lambda_i(t^i)dF_i(t^i)}{f_i(v^i)}.$$

The first term, v^i, is clearly increasing in v^i. The second term, $(1 - F)/f$, the informational rent, is typically assumed to be monotone in v^i in adverse selection models in private goods environments, by requiring the distribution to satisfy a monotone hazard rate condition.

Since the incidence of incentive taxes can have welfare effects, there is a third term to worry about, indicating that one may need more (or less!) than the standard monotone hazard rate condition to guarantee that $Q^{i\prime}(v^i) \geq 0$ is automatically satisfied when one simply plugs in virtual utilities and maximizes subject only to production feasibility. These additional conditions will imply restrictions on the distribution of welfare weights, as we illustrate in the following example.

Let v be distributed uniformly on $[0, 1]$ for all i, so $F(v) = v$ and $f(v) = 1$. Then $w(v) = 2v - \int_0^v \lambda(t)dt$ and $w' = 2 - \lambda(v)$. Therefore, the second order condition is globally satisfied for uniform distributions of valuations if and only if the maximum welfare weight is less than or equal to 2. Thus, if $\lambda(v) = 2(a + bv)/(2a + b)$, where $a \geq 0$ and $2a + b > 0$, then we are always in the "regular" case where virtual valuations are monotonic in type and the second order conditions are satisfied. If $b > 0$ (high valuation types receive more weight) then production will occur more often than in the Lindahl-Samuelson solution, while if $b < 0$, the reverse it true. However, there are λ such that the second order condition is not satisfied, even for the uniform distribution. For example, if $\lambda(v) = 3v^2$ then virtual valuations are not monotone in type; virtual valuations are decreasing for $v > \sqrt{2/3}$.

Fortunately, we can deal with these cases using a procedure called ironing. The principle behind this is flatten out the virtual valuations in the decreasing region (and often for some adjacent types as well). The trick is to choose the region is such a way that the measure of types whose virtual valuations are decreased by this flattening is balanced by types whose virtual valuations are increased. The geometry is illustrated nicely in a series of figures in Guesnerie and Laffont (1984) for the single agent case. For the example above with $\lambda(v) = 3v^2$, the optimal solution in this case will involve equal treatment of all types with valuations above $\sqrt{2/3}$; $Q^i = 1$ for all such types. Similarly, if $\lambda(v) = 3(1-v)^2$ the second order condition is violated for $v < 1 - \sqrt{2/3}$, and the optimal solution will require equal treatment of all types with sufficiently low valuations; $Q^i = 0$ for all these low- valuation types.

It is also instructive to use this example to illustrate the range of public good provision rules (or cost-benefit criteria) that are interim efficient. Suppose $N = 2$, $K = 1$. The Lindahl-Samuelson efficient outcome is to produce if and only if the average valuation exceeds 1/2, so the public good will be provided half the time.

Next suppose one shifts welfare weight to the low valuation types, to the point where $\lambda(v) = 2$ for all $v < 1/2$ and $\lambda(v) = 0$ for all $v > 1/2$. This satisfies monotonicity of virtual valuations[11] and it is easy to see that the optimal mechanism is to produce if and only if the sum of valuations exceeds 3/2. In other words, this weighting scheme effectively inflates the cost of the public good by 50 percent, so it should be produced only if the actual benefit/cost ratio exceeds 1.5. At first blush this mechanism seems like it could be improved on, since there are some states where both agents are "high" types (i.e. $v > 1/2$ for both of them), and the public good is not provided. Since all high types receive the same welfare weight, and since low types do not bear any of the cost of production in these states, it would seem to lead to an improvement in welfare. Why doesn't this lead to an improvement? The answer is that the mechanism is designed to achieve redistributive goals in addition to deciding on public good production. In this case, the welfare weights indicate that there should be a transfer from high valuation to low valuation types. Hence in the optimal mechanism there are some states where there is one low type and one high type, and the public good is not produced, but a private good transfer takes place between the low and high types. The extent of such transfers would be hindered by greater public good production due to incentive compatibility problems. We conjecture that this choice of welfare weights corresponds to the lowest possible expected output ($Q = .125$) of all interim efficient mechanisms for the uniform

[11] If one shifts the welfare weights even further downward, so that $\lambda(v) = A > 2$ for all $v < 1/A$ and $\lambda(v) = 0$ for all $v > 1/A$, then the unadjusted virtual valuations are nonmonotonic and must be ironed.. Nevertheless, from the characterization in Theorem 1, it is easy to verify that standard ironing procedures can be used and will generate an optimal mechanism with the same property: produce if and only if the sum of valuations exceeds 3/2.

case with $N = 2$ and $K = 1$.

At the other extreme, suppose the welfare weights are shifted in the opposite direction, with $\lambda(v) = 2$ for all $v > 1/2$ and $\lambda(v) = 0$ for all $v < 1/2$. In this case the optimal mechanism is to produce if and only if the sum of valuations exceeds $1/2$. In other words, the cost of the public good is effectively deflated by 50 percent, so that it should be produced if the actual benefit/cost ratio is at least .5. Again it would seem that efficiency would dictate that when both types are "low" types, the good should never be produced. However, redistributive goals implied by this welfare weighting scheme requires the low types to subsidize the cost of the public good. The most efficient way to perform this subsidization requires some "overproduction" of the public good. We conjecture that this choice of welfare weights corresponds to the highest possible expected output ($Q = .875$) of all interim efficient mechanisms for the uniform case with $N = 2$ and $K = 1$.

5 Simple Public Good Mechanisms

We next compare the efficiency of interim efficient mechanisms with the efficiency of significantly simpler mechanisms. In this section, we restrict attention to the symmetric case, where $F_i(v) = F_j(v) = F(v)$ and $\lambda_i(v) = \lambda_j(v) = \lambda(v)$ for all i, v.

5.1 Referendum mechanisms

We identify a class of particularly simple mechanisms, which uses a drastically smaller message space than the direct mechanism. In fact, each individual transmits only a single binary bit of information, which we call a "vote." Thus it is as if each individual is asked whether or not he would like to have the public good produced. If enough voters say "yes," then the public good is produced and the cost is shared equally. We call such mechanisms *referenda* with equal cost shares[12].

To be specific a *J^*-referendum* has the following three properties:

(a) Each i votes, b_i, yes ($= 1$) or no ($= 0$).

(b) The good is produced if $\sum_i b_i \geq J^*$ and is not produced if $\sum b_i < J^*$.

(c) Each i pays $\frac{K}{N}$ if it is produced and 0 if it is not.

[12] In Ledyard and Palfrey (1994) we used the term *lottery draft*, since equal cost sharing is equivalent (in expected utility) to randomly selecting, or drafting, $M \leq N$ individuals to contribute an equal (K/M) share to the production of the public good. If the private good space is discrete, randomization of this sort is needed.

Thus, in a J^*-referendum each individual casts a vote either for or against the production of the public good, which is produced if and only if at least J^* "yes" votes are cast, and costs are split equally. For each voter, it is a *dominant strategy* to vote yes, if and only if $v^i \geq K/N$. The incentive compatible direct revelation version of this is:

$$q(v) = 1 \text{ iff } \#\left\{i | v^i \geq \frac{K}{N}\right\} \geq J^*$$

$$a_i^{VL}(v) = \frac{K}{N} q(v) \; \forall i$$

The reason for considering such mechanisms is that, as we show below, they are almost interim efficient in large populations. By this, we mean that the efficiency loss from using a referendum instead of an optimal mechanism approaches zero in large populations. The two extreme referenda, corresponding to $J^* = 0$ (always produce) and $J^* = 1$ (never produce), are of independent interest and we refer to these as *command mechanisms.*

5.2 Approximate Optimality of Referenda

It is fairly easy to see that in finite populations referenda are generally interim inefficient, except in extreme cases where the critical level of J^* is equal to either 0 or N, in which case production is independent of the realization of the type-profile, v.[13] In spite of the inefficiency of the J^*-referendum, one can obtain an approximate efficiency result when N is sufficiently large. In letting N grow, we permit K to vary with N, but keep k fixed, where $k = K(N)/N$. That is, the per capita production costs of the public good are held fixed.

5.3 Per Capita Welfare Losses from Referenda

For any given set of welfare weights, λ, consider the J^*-referendum with the property that the expected sum of virtual utilities, if exactly J^* individuals vote for production of the public good, is equal to kN. For this voting rule, asymptotically in N, the public good will be produced if and only if the average virtual utility is greater than or equal to k. By the law of large numbers, this will therefore almost surely produce the optimal level of public good. (Either full production or zero production depending on whether average virtual utility exceeds or falls short of k.) Also, since the interim expected public good production ($Q_i(v^i)$) is type independent in the limit, incentive compatibility requires that the interim-expected optimal taxes approach equal cost shares as the number of agents goes to infinity. Therefore, in the limit as the number of agents goes to infinity, the J^* referendum generates the same per-capita expected welfare as the optimal mechanism.

[13]An example of this arises when v^i is distributed on the $[1, 2]$ interval for all i, and $\frac{K}{N} < 1$. In this special case, production is always optimal independent of the actual draws of v. Of course, in this case, there is no need to elicit messages from the agents at all. So $J^* = 0$ is efficient.

If there is a per-capita cost to operating a mechanism that is increasing in the size of the message space, then for a sufficiently large number of agents voting rules outperform the "optimal" mechanism computed in the previous section of this paper. This is demonstrated formally below.

Consider a sequence of J_N^*-referenda where $J_N^* = j^*N$ is set[14] such that the expected total virtual utility, if exactly j^* fraction of individuals vote "yes," equals kN. Denoting $w^+ = E[w; v > k]$ and $w^- = E[w; v < k]$, this requires choosing j^* so that $j^*w^+ + (1 - j^*)w^- = k$. What we do below is to replicate the economy, keeping the distribution of individual types constant and also keeping the per capita cost of producing the public good constant, and compare the per capita surplus using this j^* rule to the per capita surplus using the optimal rule, and show that in the limit they are the same.

Theorem 2 *Let $K_N = kN$, k fixed. Let $\lambda_i(v^i) = \lambda(v^i)$ and $f_i(v^i) = f(v^i)\forall i$. Let j^* satisfy $j^*w^+ + (1-j^*)w^- = k$. As $N \longrightarrow \infty$ the referendum mechanism using $J_N^* = j^*N$ is almost interim-efficient in the sense that it satisfies (I) and (F) and*

$$\frac{1}{N}\sum_i \int_{\underline{v}^i}^{\overline{v}^i} \lambda(v^i)u^i(\eta_N^{VL}, v^i)dF(v^i) \to \frac{1}{N}\sum_i \int_{\underline{v}^i}^{\overline{v}^i} \lambda(v^i)u^i(\eta_N^{o}, v^i)dF(v^i)$$

where η_N^{VL} denotes the J_N^*-referendum mechanism with N individuals and η_N^o denotes the optimal mechanism with N individuals.

Proof: Denote by "Y," the number of yes votes. By construction of j^*, $E[\sum_i w_i/N \mid$ "Y"$= j^*N] = k$, so that if there are at least j^*N votes, then the expected sum of virtual benefits is greater than or equal to K_N. As $N \to \infty$, by the strong law of large numbers, the expected average virtual benefit when exactly j^* fraction of the voters vote "Yes" will converge in probability to k. In other words, the probability that this J_N^* rule and the optimal rule make different production decisions for the same profile of types approaches 0 in the limit.

Now consider the reduced forms for the j^*-referendum mechanism:

$$Q_N^{VL}(v^i) = \int_{V_{-i}} q_N^{VL}(v)dF_{-i}(v) = Prob\left\{\sum_{j\neq i} b^j(V^j)/N \geq j^* - \left(b^i(v^i)/N\right)\right\}$$

where

$$b^j(V^j) = \begin{array}{ll} 1 & \text{if } v^i \geq k \\ 0 & \text{if } v^i < k, \end{array}$$

[14]Since N is finite, there is generally no exact value of j^* satisfying this equality condition. What we mean precisely is that $((J_N^* - 1)/N)E[w; v > k] + ((N - J_N^* + 1)/N)E[w; v < k] \leq k$ and $((J_N^* + 1)/N)E[w; v > k] + ((N - J_N^* - 1)/N)E[w; v < k] \geq k$.

$$\alpha_N^{VL}(v_i) = \int_{V_{-i}} a_N^{VL}(v)dF_{-i}(v) = kQ_N^{VL}(v^i)$$

and the interim-efficient mechanism

$$\begin{aligned} Q_N^o(v^i) &= Prob\left\{\textstyle\sum_{j\neq i} w^j(v^j)/N \geq k - (w^i(v^i)/N)\right\}; \text{ and} \\ A_N^o(v^i) &= \textstyle\int_{V_{-i}} a_i^o(v)dF(v). \end{aligned}$$

For large N, $Q_N^{VL}(v^i) \approx Q_N^o(v^i)$ for all v^i. Incentive compatibility then implies that for large N, $A_N^{VL}(v^i) \approx A_N^o(v^i)$ for all v^i. Therefore, for large N,

$$\int_{V^i} \lambda(v^i)u^i(\eta_N^{VL}, v^i)dF(v^i) \to \int_{V^i} \lambda(v^i)u^i(\eta_N^o, v^i)dF(v^i)$$

Since all individuals are identical, this implies

$$\frac{1}{N}\sum_i \int_{V^i} \lambda(v^i)u^i(\eta_N^{VL}, v^i)dF(v^i) \to \int_{V^i} \lambda(v^i)u^i(\eta_N^o, v^i)dF(v^i). \ QED.$$

While this is a strong result, as far as justifying the use of simple dominant strategy mechanisms for public good decisions, it would be even nicer to have a stronger result. The reason to look for a stronger result is simple. One can show that in the limit, for any optimal public good mechanism, the limit of Q_N is either 0 or 1, depending on the distribution of types and the welfare weights. Thus, using a similar argument as in the proof of the theorem above, one can show that any sequence of voting rules (or any sequence of mechanisms in general), with the property that the expected production of the public good in the limit is the same as the optimal mechanism (either 0 or 1, respectively), will also generate the same per capita welfare benefits as the optimal mechanism.

Suppose for example that $E[w] > k$. Then the command mechanism "always produce," while being suboptimal for any finite value of N (and worse than the best referendum, as well) generates the same per capita surplus as the optimal mechanism in the limit, since there is almost surely production of $q = 1$ in the limit. Moreover, any j^*N-referendum that fixes j^* less than some critical level, is asymptotically optimal. Alternately, suppose that $E[w] < k$. Then the mechanism "never produce," while being suboptimal for any finite value of N, generates the same per capita surplus as the optimal mechanism in the limit, since there is almost surely zero production in the limit. Obviously, this indeterminacy problem arises because the planner can compute the optimal limiting production decision ex ante. We next show that this indeterminacy of optimal referenda in the limit is no longer problematic if one considers a more realistic model in which the optimal limiting production decision is not known by the planner ex ante. We obtain below the much stronger result that, if there is ex ante uncertainty about the optimal limiting production decision, then there exists a *uniquely* optimal referendum rule, which is not a command mechanism.

5.4 Common Values: Uniqueness of Approximately Optimal Referendum

A simplifying assumption has been made in the above analysis, which is that the distribution of valuations is known by the planner. This simplification is convenient, but it is the source of the unsatisfying result in the previous subsection, that there is no uncertainty in the limit, so that *any* referendum – including a command mechanism – that leads to expected production that is the same as the (certain) limit of production under the optimal mechanism. We now show that a natural generalization of the model demonstrates that for a given welfare weighting scheme there is generally a *unique* referendum rule that is approximately optimal in the limit.

Define c to be the *common value component* which is added to everyone's private valuation, v_i, to generate an *adjusted individual valuation* denoted $r_i = v_i + c$. Thus c is a parameter which shifts the distribution of valuations either up or down, depending on whether c is positive or negative, respectively. We continue to use the notation as before, where $F(v)$ denotes the cdf of private valuations evaluated at v, and $\lambda(v)$ denotes the welfare weight assigned to a v-type. The common value component c is distributed according the continuous cdf, $G(c)$. One can easily show that for sufficiently low values of c, the public good should never be produced, and for sufficiently high values of c, the public good should always be produced. We assume that the support of G is bounded, but large enough to include both these very high and very low values of c, where command mechanisms are optimal. Each individual observes r_i, and the planner only knows F[15].

We now consider optimal referenda under this alternative preference structure, with this common value uncertainty.[16] We compare the per capita surplus under an optimal referendum to the per capita surplus the planner could achieve if the planner actually knew c, which is simply given by the cost-benefit criteria applied to the shifted distribution of valuations. This is clearly an upper bound on what an optimal mechanism could achieve when the planner does not know c. We show below that there exists a *unique* referendum that achieves this upper bound in the limit.

Fix a weighting scheme, λ. Then , it is easy to see that, conditional on c, the optimal mechanism requires production if and only if the average virtual valuation exceeds $k - c$. That is, the cost-benefit criterion is simply shifted by the common value parameter. Denote by c^* the critical value of c such that $E[w(v); \lambda] = k - c^*$. That is, if $c > c^*$ then the public good should always be produced in the limit, and if $c < c^*$ it should never be produced in the limit. By the earlier assumption that G has a broad support, c^* is in the interior of the support of G. Now consider the referendum defined uniquely [17]

[15] Neither the planner nor the individuals have to know G.

[16] Since we are only considering behavior under dominant strategy mechanisms here, it is not actually necessary to specify any particular common knowledge structure for the players, beyond the fact that r_i is known by i. But for concreteness, let c be common knowledge among the individuals and assume v_i is private information as before, with common prior beliefs about v given by F. When we refer to "the optimal mechanism" below, we mean optimal relative to this information structure.

[17] Moreover, j^* must be strictly between 0 and 1, since c^* is in the interior of the support of G, so the

by $E[\sum_i w_i/N \mid \#\{i : v_i + c^* \geq k\} = j^*] = k - c^*$. Voters have a dominant strategy to vote yes if and only if $v_i + c > k$. This referendum is constructed so that, in the event $c = c^*$, the expected average virtual valuation, conditional on the proportion of yes voters exactly equaling the referendum threshold j^*, is equal to $k - c^*$. Thus, if $c < c^*$ this referendum will always fail to produce the public good in the limit, while if $c > c^*$ the referendum will always produce the public good in the limit. By the argument of the previous subsection, the expected per capita welfare losses conditional on c go to 0 for every value of c (in fact, uniformly in c), except possibly at c^*. Therefore the expected per capita losses (integrating over $G(c)$) go to 0 for this referendum.

Are other referenda optimal for this weighting scheme? No. For any j-referendum with $j > j^*$, there exists a critical common value, $c^{*}* > c^*$, such that public good is only produced in the limit for realizations of $c > c^{*}*$. Thus, for realizations of c between c^* and $c^{*}*$, the public good will inefficiently fail to be produced in the limit, even though the c-conditional optimal mechanism requires production in the limit with probability approaching 1. This generates a strictly positive per capita welfare loss, conditional on c, on the order of $c - c^*$ for realizations of c between c^* and $c^{*}*$. Thus, integrating over $G(c)$, the expected per capita welfare loss is strictly positive in the limit for any j-referendum with $j < j^*$. A similar argument applies for any j- referendum with $j < j^*$.

5.5 Total Welfare Losses from j^*N-Referenda

A stronger criterion for asymptotic efficiency is the total (as opposed to per capita) surplus loss of the j^*N-referendum compared to the optimal mechanism. We prove below that the total surplus loss from using the optimal referendum instead of the optimal virtual cost-benefit criterion goes to zero in the limit. We prove this for the pure private values model (without the common value shift parameter), but it is a straightforward exercise to demonstrate that the same result holds for the more general model with the shift parameter. Moreover, for reasons shown in the previous section, the more general result is uniquely true for the referendum rule defined by $E[\sum_i w_i/N \mid \#\{i : v_i + c^* \geq k\} = j^*] = k - c^*$.

By symmetry, the total expected welfare from the optimal mechanism is equal to:

$$W_N^o = N \int_{\underline{v}}^{\bar{v}} \lambda(v)[vQ_N^o(v) - A_N^o(v)]dF(v).$$

and the expected welfare from a j^*N-referendum is:

$$W_N^{j^*} = N \int_{\underline{v}}^{\bar{v}} \lambda(v)(v - k)Q_N^{j^*}(v)dF(v).$$

Therefore, the difference in the expected total welfare (i.e., the expected welfare loss) is equal to:

$$\Delta W_N^{j^*} \equiv W_N^o - W_N^{j^*}$$

j^*-referendum is not a command mechanism.

Theorem: 2 Let $K_N = kN$, k fixed. Let $\lambda_i(v^i) = \lambda(v^i)$ and $F_i(v^i) = F(v^i), \forall i$. Let j^* satisfy $j^*w^+ + (1-j^*)w^- = k$. Then the j^*N- referendum is asymptotically interim-efficient in the sense that it satisfies (I) and (F) and:

$$\lim_{N\to\infty} \Delta W_N^{j^*} \to 0.$$

Proof: From above,

$$\Delta W_N^{j^*} = N\int_{\underline{v}}^{\bar{v}} \lambda(v)(v-k)(Q_N^o(v) - Q_N^{j^*}(v))dF(v)$$

$$-N\int_{\underline{v}}^{\bar{v}} \lambda(v)(A_N^o(v) - kQ_N^o(v))dF(v)$$

Thus, the expected welfare loss is divided into two terms. The magnitude of the first term is on the order of N times the average expected differences in the reduced form production decisions, Q_N^o and $Q_N^{j^*}$. The magnitude of the second term is on the order of N times the expected difference between equal cost sharing in the referendum and incentive compatible cost sharing in the optimal mechanism. We apply a central limit theorem below to show that both of these converge to 0 in N, although we find that convergence occurs at an order of magnitude faster rate for the second term than the first.

We begin by considering the first term, $N\int_{\underline{v}}^{\bar{v}} \lambda(v)(v-k)(Q_N^o(v) - Q_N^{j^*}(v))\, dF(v)$. Recall that both Q_N^o and $Q_N^{j^*}$ are deterministic in the limit (i.e., equal either 0 or 1). Thus if j^* is not chosen so that $Q_N^{j^*} \approx Q_N^o$, then we know that the expected welfare loss goes to infinity. However, we know from above that for j^* satisfying $E[\sum_i w_i/N \mid \#\{i : v_i \geq k\} = j^*] = k$ we are guaranteed that $Q_N^{j^*} \approx Q_N^o$. Thus, we only need to obtain a rate of convergence to 0 for $Q_N^{j^*} - Q_N^o$ and show that this converges to 0 very fast. We show below that the speed of convergence is at least on the order of $\sqrt{N}e^{-N}$, so N times the expected difference in interim quantities converges to 0, and hence the first term goes to 0 in N.

In the optimal mechanism, the good is produced if and only if $\sum_i w_i/N \geq k$. Thus, for an individual with private value equal to v^i, the interim expected output under the optimal mechanism is simply the probability that the sum of all the other virtual valuations is greater than or equal to $Nk - w(v^i)$ which equals the probability that the sample average virtual valuation of the other players is greater than or equal to $[Nk - w(v^i)]/(N-1)$. Denoting the expected value of the virtual valuation of an individual as $\overline{w}$, we know from the Central Limit Theorem that the sample average virtual valuation of $N-1$ has an asymptotically Normal distribution with mean $\overline{w}$ and standard deviation $\sigma_w/(N-1)$, where σ_w is the standard deviation of w. Thus, we get

$$Q_N^o(v) \to 1 - \Phi\left[\frac{-(\overline{w}-k) - \left(\frac{w(v)-k}{N-1}\right)}{\sigma_w/(N-1)}\right]$$

where Φ is the cumulative of the unit Normal distribution. Similarly, we can obtain an expression for the asymptotic value of $Q_N^{j^*}(v)$. It depends only on whether or not v is greater than or less than k. Denote by $b(v)$ the vote of an individual of type v, which is equal to 1 if v is greater than or equal to k and equals 0 if v is less than k. Denote by $\overline{b}$ the ex ante probability of a yes vote (which is simply equal to $1 - F(k)$), and which also equals the expected fraction of individuals voting yes. Then by a similar argument, we get that

$$Q_N^{j^*}(v) \rightarrow 1 - \Phi\left[\frac{-(\overline{b} - j^*) - \left(\frac{b(v)-j^*}{N-1}\right)}{\sigma_b/(N-1)}\right]$$

where σ_b is the variance of b.

By construction of j^*, $\lim_{N\to\infty} Q_N^o(v) = \lim_{N\to\infty} Q_N^{j^*}(v)$. That is, $\overline{b} - j^* > 0$ if and only if $\overline{w} - k > 0$. The difference $Q_N^o(v) - Q_N^{j^*}(v)$ converges to

$$Q_N^o(v) - Q_N^{j^*}(v) \approx \frac{1}{\sqrt{2\Pi}} \int_{A_N}^{B_N} e^{-x^2/2} dx$$

where

$$A_N = -\frac{\sqrt{N}(\overline{b} - j^*)}{\sigma_b} - \frac{b(v) - j^*}{\sigma_b \sqrt{N}}$$

and

$$B_N = -\frac{\sqrt{N}(\overline{w} - k)}{\sigma_w} - \frac{w(v) - k}{\sigma_w \sqrt{N}}$$

Without loss of generality, assume that $\frac{\overline{b}-j^*}{\sigma_b} > \frac{\overline{w}-k}{\sigma_w}$, so that, for sufficiently large N, $A_N < B_N$. Then for large N,

$$Q_N^o(v) - Q_N^{j^*}(v) \approx N^{1/2}\left(\frac{\overline{b} - j^*}{\sigma_b} - \frac{\overline{w} - k}{\sigma_w}\right)\frac{1}{2\Pi} e^{-N\left(\frac{\overline{w}-k}{\sigma_w}\right)^2}$$

Therefore, the expected difference between the interim expected quantities under the optimal mechanism and the j^* mechanism, $N(Q_N^o(v) - Q_N^{j^*}(v))$, is on the order of $N^{3/2}e^{-N}$, which converges to 0 in N. This establishes that the first term of the expression for the total surplus loss goes to 0.

The second term of that expression is

$$N \int_{\underline{v}}^{\overline{v}} \lambda(v)(A_N^o(v) - kQ_N^o(v))dF(v)$$

This can be rewritten as

$$\int_{\underline{v}}^{\overline{v}} \lambda(v)[N(A_N^o(v) - \overline{A}) - Nk(Q_N^o(v) - \overline{Q})]dF(v).$$

which can be further broken down into two terms:

$$\int_{\underline{v}}^{\overline{v}} \lambda(v)N(A_N^o(v) - \overline{A})dF(v)$$

and

$$\int_{\underline{v}}^{\overline{v}} \lambda(v)Nk(Q_N^o(v) - \overline{Q})dF(v).$$

Consider the second of these terms. Because $\lambda(v)$ is bounded, we just need to show that

$$\int_{\underline{v}}^{\overline{v}} N \mid Q_N^o(v) - \overline{Q}) \mid dF(v) \rightarrow 0.$$

The expression $\mid Q_N^o(v) - \overline{Q}) \mid$ is less than or equal to $Q_N^o(\overline{v}) - Q_N^o(\underline{v})$, so we only need to show that

$$\int_{\underline{v}}^{\overline{v}} N[Q_N^o(\overline{v}) - Q_N^o(\underline{v})]dF(v) = N[Q_N^o(\overline{v}) - Q_N^o(\underline{v})] \rightarrow 0.$$

Recall that $Q_N^o(v) = prob\{\overline{w} - \frac{k}{N-1} \geq \frac{w(v)}{N-1}\}$ so, using an argument similar to the one above, the difference $Q_N^o(\overline{v}) - Q_N^o(\underline{v})$ converges to

$$Q_N^o(\overline{v}) - Q_N^o(\underline{v}) \approx \frac{1}{\sqrt{2\Pi}} \int_{A_N}^{B_N} e^{-x^2/2} dx$$

where

$$A_N = -\frac{\sqrt{N}(\overline{w} - k)}{\sigma_w} - \frac{w(\underline{v}) - k}{\sigma_w \sqrt{N}}$$

and

$$B_N = -\frac{\sqrt{N}(\overline{w} - k)}{\sigma_w} - \frac{w(\overline{v}) - k}{\sigma_w \sqrt{N}}.$$

This is on the order of $\sqrt{N}e^{-N}$, so $N[Q_N^o(\overline{v}) - Q_N^o(\underline{v})] \rightarrow 0$ as $N \rightarrow \infty$. Therefore

$$\int_{\underline{v}}^{\overline{v}} \lambda(v)Nk(Q_N^o(v) - \overline{Q})dF(v) \rightarrow 0$$

as desired. By incentive compatibility, $A' = vQ'$, and by assumption $v < \overline{v} < \infty$, so it also follows that

$$\int_{\underline{v}}^{\overline{v}} \lambda(v)N(A_N^o(v) - \overline{A})dF(v) \rightarrow 0.$$

Thus $\lim_{N \rightarrow \infty} \Delta W_N^{j^*} \rightarrow 0$. QED.

6 Conclusions

In this paper, we have characterized the interim efficient public good allocation rules in a simple Bayesian public good environment, and have explored the use of simple voting schemes to approximately implements these rules. We find that the optimal mechanism involves either more or less production of the public good depending on whether the welfare weights are shifted in the direction of types with higher or lower valuations for the public good. Thus, compared to the classical optimal level of public good provision (the "Lindahl-Samuelson" solution), there should generally be some distortion. The reason for this distortion is that unless welfare weights are perfectly neutral, efficient allocations will depend in general on both the level of public good and the incidence of taxes to finance the public good. Because of incentive compatibility, the efficient way to reduce the tax burden on low-valuation (resp: high-valuation) consumers is to reduce (resp: increase) the level of provision of the public good. In the borderline case, the first-best solution is attainable only because the welfare function is independent of distribution of the private good.

We further show that for *any* interim efficient allocation rule there exists a simple dominant-strategy referendum mechanism which approximates the efficiency of that allocation rule in large populations. In a referendum, individuals simply submit a binary message (a "vote") either for or against production of the public good. If a sufficiently large fraction of the individuals vote in favor, then the public good is provided and the costs are distributed equally in the population. Otherwise, the public good is not produced. This provides an approximate "first welfare theorem" for public goods: efficient allocation rules can be (approximately) decentralized by an appropriately chosen voting rule. Moreover, if there is a common value component to the distribution of preferences, then the optimal referendum is *unique.*

There are several directions in which it would be useful to extend these results. First, the asymptotic results were obtained under an assumption on the distribution of types that guaranteed that the solution to the optimal control problem was "regular." This allowed us to conduct the analysis using only the first order conditions. In a completely general setup, we would have to include inequality constraints that could be binding if the interim utility as a function of type were not strictly convex. We expect that the main results would still hold up, but the optimal solution would involve "pooling" of types. Since the referenda we use are an extreme form of pooling of types, the results on approximate optimality of referenda should be unaffected.

Second, we note that participation constraints were not imposed in our solution for the optimum. It is fairly easy to show that when these constraints are binding, this implies a reduction in the level of the public good, since these constraints are necessarily binding on the low valuation types (Ledyard and Palfrey 1994). It is also true that, except in uninteresting cases, these constraints will imply $Q_N \to 0$ in large populations (Ledyard and Palfrey 1994, Mailath and Postlewaite 1990). But for the case of large N, it would usually seem more realistic to assume that participation is generally obligatory to all members of the group under consideration, as we have assumed here. Related to the general issue of participation is the application of the general approach presented here to excludable public goods, an extension that we will pursue in another paper. In that case, participation constraints can be dramatically relaxed by the (no-cost) exclusion of low valuation types.

Our results about the asymptotic optimality of referenda were obtained by replicating a population with the same distribution of types. In the case where distributions differ across the population, optimal referenda might involve asymmetric cost shares, although we conjecture that referenda with equal voting weight will still be asymptotically efficient. More involved extensions, such as relaxing the assumption of independent types, considering a more general form of common values, or introducing multidimensional types, remain difficult open questions.

References

[1] d'Aspremont, C. and L.-A. Gérard-Varet (1979). "Incentives and Incomplete Information." *Journal of Public Economics*, 11:25–45.

[2] Cramton, P. and T. Palfrey (1995). "Ratifiable Mechanisms: Learning from Disagreement." *Games and Economic Behavior*, 10:255-83.

[3] Crawford, V. (1985). "Efficient and Durable Decision Rules: A Reformulation," *Econometrica*, 53:817-36.

[4] Guesnerie, R. and J.-J. Laffont (1984). "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm." *Journal of Public Economics*, 25:329-69.

[5] Holmström, B. and R. Myerson (1983). "Efficient and Durable Decision Rules with Incomplete Information." *Econometrica*, 51:1799–1819.

[6] Jackson, M. and H. Moulin (1992). "Implementing a Public Project and Distributing Its Cost." *Journal of Economic Theory*, 57:125–40.

[7] Laffont, J.-J. and E. Maskin (1982). "The Theory of Incentives: An Overview," in *Advances in Economic Theory* (ed. W. Hildenbrand), Cambridge University Press: Cambridge, 31–94.

[8] Ledyard, John O. and Thomas Palfrey (1994). "Voting and Lottery Drafts as Efficient Public Goods Mechanisms." *Review of Economic Studies*, 61:327–55.

[9] Ledyard, John O. and Thomas Palfrey (1996). "Interim Efficiency in a Public Goods Problem." *Social Science Working Paper #956*, California Institute of Technology.

[10] Mailath, George and Postlewaite, Andrew (1990). "Asymmetric Information Bargaining Problems with Many Agents." *Review of Economic Studies*, 57:351-67.

[11] Mirrlees, J. A. (1971). "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*, 38:175–208.

[12] Myerson, R. (1981). "Optimal Auction Design." *Mathematics of Operations Research*, 6:58–73.

[13] Rochet, Jean-Charles (1987). "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context." *Journal of Mathematical Economics*, 16:191–200.

[14] Rochet, Jean-Charles (1995). "Ironing, Sweeping and Multidimensional Screening." Cahier de Recherche #95-11-374, GREMAQ, Universityé des Sciences Sociales, Toulouse.

[15] Wilson, Robert (1993). "Design of Efficient Trading Procedures" In *The Double Auction Market Institutions, Theories, and Evidence*, eds. D. Friedman and J. Rust. Proceedings of the Workshop on Double Auction Markets Held June, 1991 in Santa Fe, New Mexico. Proceedings Volume XIV. Addison-Wesley.