Non-Coercion, Efficiency and Incentive Compatibility in Public Goods Decisions

John O. Ledyard[1]

1 Introduction

Based on Wicksell's principles of unanimity and voluntary consent in taxation, Lindahl (1919) proposed a market-like equilibrium for public goods economies. Foley (1970) provided a proof that Lindahl equilibrium allocations are in the core of a private ownership economy. Given those results, it would not be unreasonable to think that efficient and non-coercive public goods decisions should be possible through some process yielding Lindahl equilibria. Efficiency comes from the fact that the coalition of the whole cannot block such allocations. Non-coercion, at least at a minimal level, comes from the fact that individuals cannot block such allocations.

But, in Lindahl markets, each consumer is effectively a monopsonist in the market for her own consumption of the public good. Thus, it is unlikely that she would behave like a price taker and as a result, in practice the Lindahl allocation would not obtain. Samuelson (1954) actually went further and suggested that, in public goods economies, no decentralized process could produce an allocation that was efficient. We now know that he was too pessimistic. Nevertheless, he made a good point. In considering social decisions, we need to be aware of the incentives created for the actors in the economy. Normative considerations need to be tempered by reality.[2]

One way to incorporate reality into the discussion is through incentive compatibility constraints. This is the approach of mechanism design. Accept the fact that the system one puts in place to decide on allocations will be inhabited by purposive agents. Then choose the social decision process you want so that it is compatible with the incentives of the consumers. This is the approach I will take in this paper as I explore the trade-offs between social welfare and coercion in the presence of incentive compatibility constraints.

2 Some Basics

To begin with, I introduce notation and ideas that may be familiar to the readers.

2.1 Public goods economies

There are n consumers, each with an initial endowment of a private good, $w^i \in \Re$. Consumer i will consume x^i of the private good. There is also a public good which each i

[1]Division of Humanities and Social Sciences, California Institute of Technology, Mail Code 228-77, Pasadena, California, USA 91125. email: jledyard@caltech.edu

[2]Wicksell himself seemed to aware of this. He wrote: ``How much of this ... may be of practical use in the near future, men of affairs may decide." Wicksell (1896, p. 730)

consumes in the amount $y \in \Re$. Consumer i has a utility function $u^i(y, x^i, v^i)$. The parameter $v^i \in V^i$ is the consumer's type. There is also a production side to the economy. I keep this simple and model it with a cost function $c(y)$ which is the amount of the private good necessary to produce y. I call $e = \{c(y), v^1, ..., v^n, w^1, ..., w^n\}$ an environment and let E be the set of environments under consideration.

We will be interested in allocation-tax plans, $a = (y, t^1, ..., t^n)$. Here, t^i is the amount of the tax paid by consumer i. I will often refer to these simply as allocations. The utility that i attains in e from the allocation a is $U^i(a, v^i, w^i) = u^i(y, w^i - t^i, v^i)$. A couple of definitions will be useful.

Definition 1 *An allocation is feasible (in e) if and only if the sum of the taxes collected from all the consumers is at least equal to the cost of the public good; that is,* $\sum_{i=1}^{N} t^i \geq c(y)$.

We will let F be the set of feasible allocations for e.

Definition 2 *An allocation* a' *is efficient (in e) if and only if it is feasible and there is no other feasible allocation* a^* *such that* $U^i(a^*, v^i, w^i) \geq U^i(a', v^i, w^i)$ *for all* i *and* $U^i(a^*, v^i, w^i) > U^i(a', v^i, w^i)$ *for some i.*

2.2 Mechanisms and their performance functions

A mechanism is a process through which disparate individuals in an economy can communicate and arrive at an allocation. Market equilibria and social choice functions are examples of mechanisms. We model a mechanism as a game form[3] (M, g). Here $M = M^1 \times ... \times M^n$ where M^i is i's *language* of communication - the set of messages that i can send to others. The *outcome function* $g: M \to F$ describes the allocations that arise where $g(m)$ is the allocation that occurs if each i sends the message m^i.

Given a mechanism (g, M) and an environment $\{c(y), v^1, ..., v^n, w^1, ..., w^n\}$, we get a game $(M, \rho^1(m), ..., \rho^n(m))$ where M^i is i's strategy space and $\rho^i(m) = U^i(g(m), w^i, v^i)$ is i's payoff function. Given a game, we can consider its game-theoretic equilibria. For now we use Nash-equilibria. For the game $G = (M, \rho^1, ..., \rho^n)$, the strategy $m^* \in M$ is a Nash-equilibrium of G if and only if $\rho^i(m^*) \geq \rho^i(m^* / m^i)$ for all $m^i \in M^i$ for all i.[4]

We are ultimately interested in the allocations that result from a mechanism in an

[3]In this paper I will stick with normal form games. We could consider extensive form games to deal with iteration, etc. but normal forms will be sufficient for our purposes.

[4]The expression $(m^* / m^i) = (m^{*1}, ..., m^{*i-1}, m^i, m^{*i+1}, ..., m^{*n})$.

environment for a particular equilibrium concept.

Definition 3 *The performance function of the mechanism* $\mu = (M, g)$ *in the environments* E *for Nash-equilibrium is* $P: E \to F$ *where* $P(e) = g(m^*(e))$ *and* $m^*(e)$ *is the Nash-equilbrium of the game derived from the mechanism* μ *in the environment* e.

It should be noted that a performance function is simply an *allocation function* $a: E \to F$ which specifies an allocation for each environment in E.

3 Efficient, Non-Coercive, Incentive Compatible Allocations

In spite of Samuelson's conjecture, we now know that it is indeed possible to find mechanisms, or decentralized processes, such that the performance function of the mechanism in public goods environments produces a Lindahl equilibrium in each environment.

Theorem 1 *Hurwicz (1979a), Walker(1981)*

Let $L: E \to F$ be the Lindahl equilibrium correspondence for the environments E. That is, $L(e)$ is the set of Lindahl equilibria for $e \in E$. There exists a mechanism $\mu^* = (M^*, g^*)$ such that the performance function $P: E \to F$ of the mechanism μ^* in the environments E for Nash-equilibrium satisfies $P(e) = L(e)$ for all $e \in E$.

Thus, it would seem that, even if we take into account the incentives of the consumers, there is no conflict between efficiency and non-coercion as long as we are happy with Lindahl allocations. But there is *a priori* nothing particularly special about Lindahl equilibrium allocations. Why stop there? There are many more allocations that are both efficient and non-coercive. We know from Muench (1972) that even in large economies, the core can be much larger than the set of Lindahl equilibria. Might not these core allocations be candidates for non-coercive, efficient allocations?

I will use the concept of voluntary participation to identify allocations that are not blocked by individuals. The idea is that, in a private ownership economy, a consumer can always take her endowment and just not participate in whatever process is being used. Voluntary participation seems to me to be a necessary condition for non-coercion.[5]

Definition 4 *A feasible allocation* a *in an economy* e *satisfies voluntary participation for consumer* i *if and only if*

$$u^i(y, w^i - t^i, v^i) \geq u^i(0, w^i, v^i). \tag{1}$$

It may of course be possible for the consumer to ``not participate" by not paying taxes and still consuming the public good if it is not excludable. Or the consumer could take her resources and the technology to produce the public good when she does not participate. Each of

[5] It does assume that private property rights are enforced. We discuss this assumption below in section 7.2.

these would provide a higher value for the right hand side and a smaller set of allocations satisfying non-participation. But, I will stay with the definition we have since it is the weakest and sufficient for the result of interest. For ease in notation, from this point forward, I will normalize each i's utility so that $u^i(0, w^i, v^i) = 0$.

I want to identify all allocations that are efficient and satisfy voluntary participation. But I also want these allocations to be achievable in the sense that there is a mechanism whose Nash-equilibria will yield those allocations. Such allocations are called incentive compatible.

Definition 5 *Given a set of environments* E *we say that the allocation function* $a : E \to F$ *is Nash incentive compatible on E if and only if there is a mechanism* $\mu = (M, g)$ *such that the performance function* $P : E \to F$ *for* μ *in* E *for Nash-equilibrium satisfies* $P(e) = a(e)$ *for all* $e \in E$.

I am now equipped to state a rather remarkable theorem.

Theorem 2 *Hurwicz (1979b)*
Given a set of public goods environments E that is rich enough,[6] suppose there is an allocation function $a : E \to F$ such that a is efficient (for all $e \in E$), is Nash incentive compatible, and satisfies voluntary participation. Then $a(e) \subseteq L(e)$ for all $e \in E$.

That is, the only allocation function that is efficient, non-coercive, and Nash incentive compatible is the Lindahl equilibrium allocation function.[7]

One might wish to stop at this point since it seems that Lindahl has been validated. But, there is problem from a strict game-theoretic point of view. Nash-equilibrium is a complete information concept. It is an appropriate game theoretic equilibrium only when all of the consumers know all of the details of the environment. We are usually interested in situations where each consumer knows only their own type (w^i, v^i) and not the types of the others. For that, I need a new game theoretic model of behavior.

4 Incomplete Information

In this section I examine the tradeoffs between efficiency and non-coercion when consumers have incomplete information about the environment. Because we are now in a world of uncertainty, I will assume that consumers are von Neumann - Morgenstern decision makers who act as if they have beliefs and maximize expected utility. To model this, I add one concept

---

[6] A sufficient condition for ``rich enough" is that all CES utility functions are in E. Weaker conditions are possible.

[7] In a comment on the Hurwicz results, Thomson (1979) proves for private goods economies if one replaces voluntary participation with fairness (non-envy) then one gets that the allocation function must be Walrasian from equal endowments. I am sure that the same type of result holds in public goods economies where Walrasian is replaced with Lindahl.

to the previous complete information model; a prior distribution over possible types,[8] $\pi(v^1,...,v^n)$. To make things simpler, I will assume consumers have quasi-linear preferences: that is, $u^i(y,w^i-t^i,v^i)=\phi^i(y,v^i)+w^i-t^i$.

To make things interesting, I will assume that there is always a public goods problem. That is, I assume (i) sometimes it is efficient to produce the public good; if $a(v)$ is efficient for all $v\in V$, then $\int y(v)d\pi(v)>0$, and (ii) some types get no benefit from the public good; for each i there is $v_0^i\in V^i$ such that $u^i(y,w^i,v_0^i)\le u^i(0,w^i,v^i)=0$.

As is standard, I distinguish two different incomplete information situations: one in which consumers know their own type but know nothing else (called the interim situation) and one in which consumers know nothing, not even their own type (called the ex ante situation).

4.1 Interim Information

In the interim information condition, when the consumer knows her own type v^i, her (expected) utility for an allocation is:

$$U^i(a(\cdot)\,|\,v^i)=\int_{V^{-i}}u^i(a(v),v^i)d\pi(v\,|\,v^i) \tag{2}$$

where $V^{-i}=V^1\times...V^{i-1}\times V^{i+1}\times...\times V^n$ and $\pi(v\,|\,v^i)$ is the conditional distribution on V^{-i} given v^i.

A mechanism (M,g) and an incomplete information environment (π,E), where π is the prior beliefs and E is a set of complete information environments, combine to create a game with incomplete information, $(M,\rho^1(m,v^1),...,\rho^n(m,v^n),\pi)$. In these games, a strategy for i is a function $\beta^i:V^i\to M^i$. The relevant game theoretic equilibrium for incomplete information is Bayes Equilibria.

Definition 6 *For the game* $G=(M,\rho^1,...,\rho^n,\pi)$, *the strategy* β^* *is a Bayes Equilibrium of G if and only if for all* $v\in V, m^i\in M^i, i$,

$$\int_{V^{-i}}u^i(g(\beta^*(v)),v^i)d\pi(v\,|\,v^i)\ge\int_{V^{-i}}u^i(g(\beta^*(v/m^i)),v^i)d\pi(v\,|\,v^i). \tag{3}$$

I am interested in the allocations that arise from a mechanism. By the revelation principle, P is the performance function of some mechanism (M,g) in the environment (π,E) for Bayes equilibrium, where $P(e)=g(\beta^*(v))$, if and only if $P(\cdot)$ is interim incentive compatible in (π,E).

---

[8] We could include endowments, w^i, in the type but that just adds notation.

Definition 7 *A feasible allocation (or a performance function),* $a: v \to F$ *, is interim incentive compatible (IIC) in* (π, E) *if and only if for all* $v^i \in V^i$ *and for all* i *,*

$$U^i(a(\cdot) \mid v^i) \geq \int_{V^{-i}} u^i(a(v/v^{*i}), v^i) d\pi(v \mid v^i) \tag{4}$$

where $(v/v^i) = (v^1, ..., v^{i-1}, v^{*i}, v^{i+1}, ..., v^n)$.

Because of the incomplete information, I also need a new concept for non-coercive allocations. In the interim situation, a feasible allocation $a(\cdot)$ is (individually) non-coercive only if it satisfies an interim voluntary participation constraint.

Definition 8 *A feasible allocation* $a: V \to X \times Y$ *satisfies interim voluntary participation (IVP) if and only if* [9]

$$U^i(a(\cdot) \mid v^i) \geq 0 \text{ for all } v^i \in V^i, \text{ for all } i. \tag{5}$$

Given these concepts, I can now explore whether in the interim information situation we can find a social choice rule or mechanism which is interim incentive compatible, interim non-coercive and efficient. Unfortunately the general answer is no.

Theorem 3 *Guth-Hellwig (1986), Malath-Postlewaite (1990), Hellwig (2003)*

There is no mechanism (M, g) whose performance function in (π, E) for Bayes equilibrium is efficient, incentive compatible (IIC), and satisfies interim voluntary participation (IVP).

The idea is that with positive production in some environments required by efficiency, those unlucky enough to be very low types, near v_0, will be worse off (in iterim utility) than if they could exit, avoid taxation, and live off their endowments. There is an unavoidable conflict between non-coercion and efficiency under the interim information condition.

The conflict becomes extreme in larger economies. To see this, let's give up efficiency and just ask what non-coercive, incentive compatible mechanisms are possible. The answer is not good. In large economies if per-capita costs are not infinitesimal then any allocation that is interim incentive compatible and satisfies interim voluntary participation has the property that the probability of producing the public good is infinitesimal.

Theorem 4 *Mailath - Postlewaite (1990)*

Let $\mu^n = (M^n, g^n)$ and (E^n, π^n) be a sequence of mechanisms and sets of economies as $n \to \infty$. Let $a^n(v) = (y^n(v), t^n(v)$ be the performance function of μ^n in (E^n, π^n). Suppose that, for all n, there is a positive constant, δ such that $c^n(y^n(v)) > n\delta + \sum_i u^i(y^n(v), w^{ni} + t^{ni}(v), v^i)$ for some $v \in V^n$. Then $\int_V y^n(v) d\pi^n(v) \to 0$ as $n \to \infty$.

---

[9] Remember that I normalized utility so that $u^i(0, w^i, v^i) = 0$.

The reason is very intuitive and worth repeating. As the group grows large, if $y(v)$ is an efficient allocation for all $v \in V$, then i's report about v^i has less and less effect on the choice. Therefore by IIC, i's tax must depend less and less on i's report. In the limit this means i's tax is constant no matter what i says. If that is true for everyone, then their taxes are equal to the per-capita cost of the public good. Therefore when i has very low utility for the public good, when their value v^i is less than the per capita cost, they will be worse off than if they did not participate at all. That is, IVP can not hold unless the probability of producing the public good is very small. Thus in the limit there can be no production of the public good.[10]

This is really bad news. It says that in very large groups if we impose individual interim voluntary participation constraints, we are doomed to zero public good production. The conflict between non-coercion and efficiency is as bad as it gets in interim information situations in large economies.[11] To sidestep the interim conflict, some argue that mechanisms should not be chosen in the interim information situation but instead in the ex ante information situation. That is, one should go behind Rawl's veil of ignorance. Let us see what happens when we do that.

4.2 Ex Ante information

In the ex ante situation, are there mechanisms, or allocations, that are feasible, incentive compatible, non-coercive and efficient? We need to use the relevant information concept for each. We will use interim incentive compatibility since that is the information state when the mechanism is deployed. But I want an ex ante concept for voluntary participation.

In the ex ante information condition, when a consumer does not know her type, her (expected) utility for an allocation is:

---

[10] A formal argument can be easily made in linear environments, where $u^i(y, x^i, v^i) = v^i y + x^i, v^i \in [v_0, v_1], c(y) = nky$, and $k \in (v_0, v_1)$. Efficiency requires that $y(v) = 1$ if and only if $\frac{\sum v^i}{n} \geq k$. $y(v) = 0$ otherwise. Let $Q(v^i) =$ prob $\{\frac{\sum_{-i} v^i}{n} \geq k - \frac{v^i}{n} \mid v^i\}$ and $T(v^i) = \int_{V-i} t^i(v) d\pi(v \mid v^i)$. Then $\partial Q(v^i)/\partial v^i \to 0$ as $n \to \infty$. Interim incentive compatibility requires, assuming continuity, $v^i \partial Q(v^i)/\partial v^i - \partial T(v^i)/\partial v^i = 0, \forall v^i$. Therefore, $T(v^i) = T^i$ where by feasibility $\sum T^i = kn \int y(v) d\pi(v) = kY$. That is, incentive compatibility and feasibility combine to require per - capita taxation for all i. Assume symmetry and let $t^i(v) = k, \forall i, v$. Remember i's utility is then $U^i(a(\cdot), v^i) = (v^i - k)\text{prob}[y(v) = 1]$. For some $i, v_0 \geq v^i < k$. Thus IVP is violated unless $\text{prob}[y(v) = 1] = 0$.

[11] Hellwig (2003) provides a more optimistic result by changing the assumption on costs. He assumes that costs are independent of the size of the economy so that if the efficient level of the public good is bounded, as it is in the linear economy where $y \in [0,1]$, then per capita costs will become infinitesimal in large economies. Then the solution to $\max \int_V \sum u^i(y(v), w^i - t^i(v), v^i) - c(y(v) d\pi^n(v)$ subject to $a(\cdot)$ is IVP and IIC and will have the property that in large economies $a(\cdot)$ is approximately efficient. Of course, this is because in Hellwig's large economies the maximum amount of the good should almost always be produced.

$$U^i(a(\cdot)) = \int_V u^i(a(v), v^i) d\pi(v). \tag{6}$$

Definition 9 *A feasible allocation* $a(\cdot)$ *satisfies ex ante voluntary participation (EVP) if and only if*

$$U^i(a(\cdot)) = \int_V u^i(a(v), v^i) d\pi(v) \geq 0, \forall i \tag{7}$$

Is there a mechanism that is efficient, interim incentive compatible, and satisfies ex ante voluntary participation? Perhaps surprisingly in light of the results in the previous section, the answer is yes. Following Bierbrauer (2010), I take this in two steps.

First, D'Aspremont & Gerard-Varet (1979) and Arrow (1979) have shown us there are mechanisms whose performance functions are efficient and interim incentive compatible. These AGV mechanisms are VCG mechanisms (Vickrey (1961), Clarke (1972), Groves (1973)) using the prior beliefs in a clever way to balance the budget. They are generally referred to as expected externality mechanisms since everyone is taxed the expected externality they cause for the rest of the group through their participation.

Second, remember that, in quasi-linear environments, efficient allocations balance the budget (that is, $\sum t^i(v) = c(y(v))$ and maximize surplus (that is, $y(v) \in \arg\max_y \sum \phi^i(y, v^i)$). If, as I have assumed, when y is efficient $\int y(v) d\pi(v) > 0$, then the expected surplus of an efficient allocation function, $\int_V \max_y [\sum \phi^i(y(v), v^i) - c(y(v)] d\pi(v)$, will be positive. Take any AGV mechanism and add lump sum taxes so that every consumer shares the expected surplus equally. This gives a new mechanism which remains incentive compatible and efficient and now satisfies ex ante voluntary participation.

Proposition 1 *Bierbrauer (2010)*

There are mechanisms whose performance functions are efficient, interim incentive compatible, and satisfy ex ante voluntary participation.

That is, there is no conflict between efficiency and non-coercion in the ex ante information condition even if we impose incentive compatibility constraints. But what do these allocations look like?

4.2.1 Efficiency, non-coercion, and voting

What do ex-ante non-coercive, efficient allocations look like? Could they come from any institutions that we already know? For example, are they Lindahl equilibria? For now I leave this as an open question and, instead, look at a special case.

It turns out that for large electorates in linear public goods economies[12] these ex ante non-coercive, efficient mechanisms do look like something we know. In particular they can be approximated by q-referenda. A q-referendum begins with individuals voting yes or no on whether to produce the public good. The good is produced if the percentage of yes votes is greater than or equal to q. If the good is produced everyone pays k in taxes. If the good is not produced no one pays. Using results from Ledyard-Palfrey (2002), one can show

Theorem 5 *Let* q *satisfy* $qE[v \mid v > k] + (1-q)E[v \mid v < k] = k$.[13] *Let* a^o *be the surplus maximizing allocation and let* a^q *be the performance function of the q - referendum. Then* $\lim_{n\to\infty} \sum \int u^i(a^o(v), v^i)\, d\pi(v) - \sum \int u^i(a^q(v), v^i)\, d\pi(v) = 0.$

There is a q-referendum that generates allocations that are interim incentive compatible and, in large economies, approximately efficient. If the types are distributed symmetrically, then the q-referendum allocations also satisfies ex ante voluntary participation.

In fact, there is an even stronger result. These q-referenda are ex post incentive compatible in the sense that, even after all voters know every thing, no one wants to change their vote. q-referenda are dominant strategy mechanisms. Thus, they do not depend on the existence of a common prior or common knowledge of information and rationality. The only place that the beliefs play a role is in the determination of q.[14]

5 Future Possibilities?

In the previous section, I have relied heavily on the model of Bayes equilibrium for economies with incomplete information. Many, including myself, view the required underlying assumptions of common knowledge of information and rationality as unrealistic. Often, as an alternative, one redirects the search to try to find mechanisms in which individuals can act rationally without any common knowledge. One class of such mechanisms contains those whose games have dominant strategy equilibria.

Definition 10 *For the game* $G = (M, \rho^1, ..., \rho^n)$*, the strategy* β^**, where* $\beta^i : V^i \to M^i$, is a dominant strategy equilibrium of G if and only if for all $v \in V, m^i \in M^i, i,$

$$u^i(g(\beta^*(v)), v^i) \geq u^i(g(\beta^*(v)/m^i), v^i). \tag{8}$$

---

[12]Linear economies have $u^i = v^i y + w^i - t^i, c(y) = kny, y \in \{0,1\}, \pi$ has full support on $V = [v_0, v_1]^n$, and $v_0 < k < v_1$.

[13]If $k = (E[v \mid v > k] + E[v \mid v < k])/2$ then $q = 1/2$.

[14]I must admit this result is not so interesting for independent values since then, as $n \to \infty$, it becomes certain whether to always produce the public good or never produce the good so $q = 0$ or 1. But if values are correlated this becomes a lot less trivial.

I am particularly interested in the allocations that arise from mechanisms with dominant strategy equilibria. By the revelation principle, $P: V \to F$ is the performance function of such a mechanism if and only if P is ex post incentive compatible.

Definition 11 *A feasible allocation* $a(\cdot)$ *is ex post incentive compatible (EIC) in* E *if and only it for all* $v \in V, m^i \in V^i, i$

$$u^i(a(v), v^i) \geq u^i(a(v/m^i), v^i). \qquad (9)$$

This is called ex post incentive compatible because it is an equilibrium of the game played in the ex post information situation when all information is known. It should be noted that ex post incentive compatibility is stronger than interim incentive compatibility. If a mechanism is ex post incentive compatible then it is also interim incentive compatible but not the reverse. Further, unlike the interim approach, ex post incentive compatibility requires neither common knowledge of information nor common knowledge of rationality.

Are there allocations that are efficient, ex post incentive compatible, and satisfy voluntary participation? It is well known that, no matter what type of voluntary participation constraint is considered, the answer is no.

Theorem 6 *If the class of environments* E *is rich enough, there is no mechanism* (M, g) *whose performance function on* E *in efficient and ex post incentive compatible.*

Does this mean we need to give up our search for allocations that are efficient, non-coercive and ex post incentive compatible? Maybe not. In a recent paper, Krajbich et al (2009) suggest that future technology may be able to significantly change the incentive compatibility constraints we have been using. Based on research in neuro-economics, they first show for a very special case that it is possible to observe a neuro-signal using MRI that is correlated with the consumer's value of the public good. Then they show that if the tax paid by consumers depends appropriately on both the consumer's claimed value for the public good and the signal, the consumer will have an incentive to correctly report their true value for the good. This possibility has some interesting implications for our discussion of efficiency and non-coercion.

The signal technology can be represented by a conditional probability function where $f(s, v)$ is the probability that we will observe the signal s if the agent's true type is v. Using the revelation principle, we need only consider direct mechanisms, where $M^i = V$, augmented with the signal so that $h(m, s) = [y(v), t(v) + r(v, s)]$. The game-theoretic equilibrium we will use is that of dominant strategies in m before the observation of s.[15]

Definition 12 *The augmented mechanism* (V, h) *is ex post incentive compatible in* E *if and only if for all* $v \in V, m^i \in V^i$, *and for all* i,

[15]This timing is crucial to all of the discussion that follows. I discuss this in more detail in section 7.1 below.

$$\phi^i(y(v),v^i)+w^i-t^i(v)-\int_S r^i(v,s)dF(s,v)\geq \tag{10}$$
$$\phi^i(y(v/m^i),v^i)+w^i-t^i(v/m^i)-\int_S r^i(v/m^i,s)dF(s,v). \tag{11}$$

Definition 13 *The augmented mechanism* (V,h) *satisfies ex post voluntary participation in* E *if for all* $v\in V$, *and for all* i,

$$\phi^i(y(v),v^i)+w^i-t^i(v)-\int_S r^i(v,s)dF(s,v)\geq 0 \tag{12}$$

Krajbich et. al. are able to establish that if the signal technology satisfies a condition originally identified by Cremer and McLean (1985, 1988)[16] then

Theorem 7 *Krajbich et. al. (2010)*

Let $a(\cdot)$ be an allocation function that is efficient and satisfies voluntary participation ($u^i(a(v),v^i)\geq 0,\forall v$). Then there is an augmented mechanism that is ex post incentive compatible and satisfies ex post voluntary participation, and whose performance function, $a^a(\cdot)$, yields the same expected outcome.[17]

With the availability of a signal technology, even if we require dominant strategies, incentive compatibility imposes no constraints on the choice of efficient and non-coercive allocations. In particular there is an ex post incentive compatible mechanism satisfying ex post voluntary participation whose performance function is the Lindahl equilibrium allocation function.

6 Summary to here

My goal in this paper is to explore options for efficient and non-coercive public good decisions when the choices are constrained to be incentive compatible. We have seen that the answers depend crucially on which concept of incentive compatibility I use. This, in turn, depends on which behavior model of consumers I use and what I assume about their state of information.

In a world of complete information where consumers are price takers, Lindahl equilibrium allocations are an option. Further, if consumers are more strategic and act in accord with Nash-equilibrium, Lindahl equilibrium allocations become the only option.

In a world of incomplete information, with Bayes-equilibrium behavior, the options

---

[16]The condition is that, for all $v\in V$, the vector $f(\cdot,v)\in\Delta(S)$ is not in the interior of the convex hull of all such vectors.

[17]That is, $y^a(v)=y(v)$ and $t^a(v)=t^a(v)+\int_S r^a(v,s)dF(s,v)$.

depend on the state of information. In the interim information situation, there is an impossibility theorem. There are no allocations which are efficient, non-coercive and incentive compatible. Worse yet, in large economies, the only allocations which are non-coercive and incentive compatible involve virtually no production of the public good.

In the ex ante information situation, there is an existence theorem. There are efficient, non-coercive and incentive compatible allocations. Further, tor the very special case of linear public goods economies, these allocations are approximated by voting mechanisms called q-referenda.

Finally, in a world with a technology generating signals correlated with individual values, even if one asks for mechanisms that have dominant strategies, incentive compatibility puts absolutely no constraint on our choice of efficient and non-coercive allocations so long as decisions are made before the signals are generated.

But we are still left with a number of open questions. In the rest of this paper, I want to mention two of these.

7 Commitment and Enforcement

7.1 Timing

It is tempting to think of the ex ante and interim situations not as information conditions but as a sequence in time. In this view, decisions are first made in the ex ante stage and then they are played out, with perhaps more decisions, in the interim stage. If one thinks like this, it raises a host of new questions, most of which remain open.

First, if the mechanism chosen in the ex ante stage is efficient, interim incentive compatible, and satisfies ex ante voluntary participation, then by theorem 3 there is a positive probability that at the interim stage some consumer will not satisfy the interim voluntary participation constraint.[18] That is, they will want to take their endowments and not participate.

Suppose , at the ex ante stage, someone anticipates the interim stage and proposes that decisions be made subject to interim voluntary participation. This would, of course, mean giving up efficiency, but one could still maximize expected surplus subject to the constraints. The problem is,

Theorem 8 *Bierbrauer (2003) Let* $a(\cdot)$ *be an allocation that is interim incentive compatible and satisfies interim voluntary provision. Then, at the ex ante stage, there is an efficient, interim incentive compatible allocation satisfying ex ante voluntary participation,* $\hat{a}(\cdot)$ *such that*

$$U^i(\hat{a}(\cdot)) > U^i(a(\cdot)) \forall i. \tag{13}$$

[18]By theorem 4, as $n \rightarrow \infty$ this probability becomes a certainty.

Any proposal to impose interim voluntary participation at the ex ante stage would be defeated unanimously. The issue is commitment. With commitment on the part of everyone not to defect when and if the interim voluntary participation constraint is violated at the interim stage, implementing the efficient, incentive compatible allocation satisfying voluntary participation goes smoothly. If there is no such possible commitment, then from a game theoretic point of view, we should introduce something like a sub-game perfection constraint at the ex ante stage. This would certainly involve interim voluntary participation. But undoubtedly there is more. For example we may want to rule out choosing mechanisms in the ex ante stage that will be unanimously voted out in the interim stage. A minimal requirement for this would be that the mechanisms be interim incentive efficient.[19] What else is involved in sub-game perfection for mechanisms remains an open question.

Notice that the same issue arises if we introduce a third stage, generally called the ex post stage, at the time when everyone knows everything. Then, even if interim voluntary participation is satisfied, ex post voluntary participation may not be. That is, it is possible that $u^i(a(v), v^i) < u^i(0, w^i, v^i)$. This happens, for example, if one is on the losing side of a referendum.

The same issue is involved in the mechanisms of section 5 where decisions are made before observing the signal. There is no guarantee that consumers would, given the opportunity, be willing to voluntarily participate after their signal is known. That is because it is possible that $u^i(y(v), w^i - t^i(v) - r^i(v,s), v^i) < u^i(0, w^i, v^i)$. The consumer is willing to make the bet on participation before the signal is known but may be unhappy and regret the decision after. This is true even if the consumer is risk-neutral as I have assumed above. It is also true that if the consumer can report their value after the signals are known, then there is no mechanism for which it will be a dominant strategy to report truthfully. With a commitment to live with the result of the signal, we can get efficiency, incentive compatibility and voluntary participation. Without that commitment, we do not.

We really need a better model that takes into account the timing and repeated nature of the public goods decision problem. But that might be good news. Up to now we have been considering a single (perhaps multi-dimensional) decision along with a sequence of times to get to that decision and beyond it. But what if the group of consumers is going to be confronted by a series of public good decisions over time then sometimes an individual would want the public good, ${}^1v^i > {}^1k$, and sometimes they may not, ${}^2v^i < {}^2k$. Taken over a long time, this is like being in the ex ante stage. Losing one election is not so bad if, on average, the winning compensates for the losing. But, again, the possible problem is the inability of a group to commit. Anyone who has been involved in faculty decisions over time knows how hard it is to enforce inter-temporal agreements. Even the US Congress has this problem. Some form of sub-game perfection will need to be incorporated into incentive incompatibility to deal with this.

---

[19] That is, there is no other mechanism for which it is common knowledge - at the interim stage - that everyone would be at least as well off and some would be better off.

7.2 The Guardians

One question always lurking in the background of any mechanism design paper is ``But who will guard the guardians?"[20] There is always, at some point in the analysis, a reliance on explicit or implicit enforcement of the rules of the game. The voluntary participation constraints I have been using rely heavily on the explicit commitment to enforce property rights. If endowments can be confiscated, then voluntary participation constraints can be ignored. The incentive compatibility constraints I have been using rely heavily on the implicit commitment of the mechanism to actually implement the public good levels and taxes required by the reports of the consumers. If those implementing the rules can change their minds after seeing the reports of the consumers, then our positive analysis is wrong. Anticipating the lack of commitment, consumers will behave differently than we have modeled. For example, in the augmented mechanism of section 5 our dominant strategy model predicts consumers will report their true value when asked. But, if they anticipate that the managers of the mechanism process will do something other than advertised, the consumers would be rational to report other than truthfully.

In all of our analysis, there is an intended game form and there is, what Hurwicz called, a true game. Is the intended game and its equilibria self-enforcing in the context of the true game? How does this affect the revelation principle? I have some thoughts on these and other relevant questions. But that is for a future paper.

[20] This question from the Roman author Juvenal, is the title of Leo Hurwicz's Nobel Prize Lecture (Hurwicz 2007).

Bibliography

Arrow, K. (1979) ``The property rights doctrine and demand revelation under incomplete information" in Boskin, M. editor, Economics and Human Welfare, Academic Press, New York.

Bierbrauer, F. (2009) ``On the legitimacy of coercion for the financing of public goods"" Preprints of the Max Planck Institute for Research on Collective Goods, Bonn 2009/15

Clarke, E. (1972) "Multi-part Pricing of Public Goods" Public Choice 11: 17-33

Cremer, J. & McLean, R.P. (1985) ``Optimal Selling Strategies Under Uncertainty for a Discriminating Monopolist When Demands are Interdependent" Econometrica 53, 345-361

Cremer, J. & McLean, R.P. (1988) ``Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions" Econometrica 56, 1247-1257

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

Foley, D. (1970) ``Lindahl's Solution and the Core of an Economy with Public Goods " Econometrica, Vol. 38, No. 1, pp. 66-72

Groves, T. (1973) "Incentives in Teams" Econometrica 41: 617-631

Güth, W. and M. Hellwig (1986). " The Private Supply of a Public Good." *Journal of Economics*, 5: 121-59.

Hellwig, M. (2003). " Public Good Provision with Many Participants." *Review of Economic Studies*. 70:589-614.

Hurwicz , L. (1979a) ``Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash-equilibrium Points" The Review of Economic Studies, Vol. 46, No. 2, pp. 217-224

Hurwicz, L. (1979b) ``On allocations attainable through Nash-equilibria" in Laffont, J-J editor, Aggregation and Revelation of Preferences, North-Holland,

Hurwicz, L. (2007) "But who will guard the guardians?" Nobel Prize Lecture, *http://nobelprize.org/nobel_prizes/economics/laureates/*2007*/hurwicz−lecture.html*

Krajbich, I., Camerer, C.F., Ledyard, J. & Rangel, A. (2009) ``Using neural measures of economic value to solve the public goods free-rider problemÓ Science 326, 596-599

Krajbich, I., Ledyard, J., Camerer, C.F., & Rangel, A. (2010) "Neurometrically Informed Mechanism Design", working paper, Caltech, Pasadena, CA

Ledyard, J. and T. Palfrey (2002). " The Approximation of Efficient Public Good Mechanisms by Simple Voting Schemes." *Journal of Public Economics* 83: 153-72..

Lindahl, E. (1919/1967) ``Die Gerechtigkeit der Besteurung, translated (in part) as ``Just Taxation: A positive solution" in R. Musgrave and A. Peacock, editors, Classics in the Theory of Public Finance, Macmillan, pp. 168-176

Mailath, G. and A. Postlewaite. (1990). " Asymmetric Information Bargaining Problems with Many Agents." *Review of Economic Studies*. 57:351-67.

Muench, T., (1972) "The core and the Lindahl equilibrium of an economy with a public good: an example," Journal of Economic Theory, Elsevier, vol. 4(2), pages 241-255, April.

Samuelson, P. (1954) ``The theory of public expenditure" Review of Economic Studies 36, pp. 387-389

Thomson, W. (1979) ``Comment" in in Laffont, J-J editor, Aggregation and Revelation of Preferences, North-Holland,

Vickrey, W. (1961) "Counterspeculation, Auctions, and Competitive Sealed Tenders" Journal of Finance 16: 8-37.

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

Wicksell, K. (1896) ``A New Principle of Just Taxation" translated by J. Buchannan, in R. Musgrave and A. Peacock editors Classics in the Theory of Public Finance, Macmillan