forthcoming in *International Review of Law and Economics*

The Provision of Environmental Protection Measures under Incomplete Information: An Introduction to the Theory of Mechanism Design

Winand Emons*

Volkswirtschaftliches Institut, Abteilung für Wirtschaftstheorie, Universität Bern, Gesellschaftsstraße 49, CH-3012 Bern, Switzerland

ABSTRACT

In this paper we use several results from the mechanism design literature to argue that the efficient provision of environmental protection measures by voluntary Coasian bargaining procedures is generally impossible. Voluntary institutional arrangements typically lead to an undersupply of environmental protection. If our aim is to provide the efficient amount of environmental protection, we cannot but rely on institutions that have the power to use collective coercion.

Keywords: environmental protection, public goods, incomplete information, mechanism design.

Journal of Economic Literature Classification Numbers: D82, H41.

*I thank Aleks Berentsen, Katia Delbiaggio, Christoph Eisenring, Michael Faure, Martin Hellwig, Anthony Ogus, Göran Skogh, and two referees for helpful comments.

I. Introduction

In this paper we argue that the efficient provision of environmental protection measures by voluntary bargaining procedures is generally impossible. Voluntary agreements typically lead to an undersupply of environmental protection. If the aim is to provide the efficient amount of environmental protection, we cannot but rely on institutions that have the power to use collective coercion.

Our basic premise is the observation that the majority of environmental protection measures are public rather than private goods. Since several agents benefit from any environmental protection measure, we have to compare the measure's costs with the consumers' aggregate benefits in order to reach a sensible decision concerning the provision of the public good. If each agent's benefit from the measure were commonly known, the decision about the provision would be straightforward: aggregate the individual benefits and go ahead with the measure if aggregate benefits exceed costs. Yet, as was stressed by Samuelson (1954), it is unreasonable to assume that all consumers' valuations are commonly known. The more realistic case to consider is the one in which each beneficiary knows how much the environmental protection measure is worth to herself, but is incompletely informed about the value of the public good to the other consumers.

In such a situation of incomplete information the crucial question arises what incentives an agent has to reveal her information so that the right decision concerning the provision of the public good can be taken? More specifically, can we design institutions under which consumers reveal how much the environmental protection measure is worth to them, thus providing the means for the efficient decision to be taken? The answer to this question is known as the problem of mechanism design for the efficient provision of public goods. In this paper we want to convey the basic ideas of this important, though fairly technical, mechanism design literature by means of a simple example.

We first describe two mechanisms permitting the efficient provision of environmental protection. Under these mechanisms the agents truthfully reveal how much the public good is worth to them thus making the efficient decision possible. Yet, these two mechanisms reach this desired feature at a cost — some individuals must be forced to participate. That is, some individuals do worse if they participate in the mechanisms than if they do not take part in the schemes which, in turn, implies that the status quo without the public good is retained. These individuals will not voluntarily participate in the mechanisms; accordingly, for the mechanisms to work they must be forced to participate. These mechanisms thus rely on collective coercion.

The next question we ask, therefore, is whether there are mechanisms permitting both, the efficient provision of public goods and the voluntary participation of agents? The answer to this question is, unfortunately, negative. If we insist on voluntary participation, there will typically be an undersupply of environmental protection measures. This important negative result implies that voluntary Coasian bargaining arrangements lead to an inefficient undersupply of public goods. We even strengthen this negative result in the following sense: the inefficiency of mechanisms in which the agents voluntarily participate increases with the number of beneficiaries. The larger the number of beneficiaries of an environmental protection measure, the larger is the discrepancy between the efficient amount of the public good and the amount actually provided.

Consequently, on the one hand, if we want institutions in which everybody voluntarily participates, we will have an undersupply of public goods — viz. the current state of the environment. On the other hand, if we strive for efficiency, we must use institutional arrangements relying on collective coercion — viz. the current difficulties in implementing the desired environmental protection measures.

The remainder of this paper is organized as follows. Section 2 outlines the example and derives the efficient decision rule for the provision of the public good. The next two sections describe the incentive problem and a helpful tool for simplifying the analysis. In sections 5 and 6 we derive the two efficient mechanisms. The following two sections describe the impossibility results. Section 9 concludes the paper.

II. Efficient Provision of a Discrete Public Good

Throughout the paper we will consider the following example. A social planner considers building a sewage treatment plant. The plant will reduce the offensive smells of a river. Consumers living downstream of the plant will enjoy fewer offensive smells or, in short, better air quality. The decision whether or not to build the plant turns out to be problematic because air quality is a *pure public good*: First, air quality is *nonrival*, i.e., one person's consumption does not reduce the amount available to the other consumers. Second, air quality is *nonexcludable*. It is not possible to exclude a certain downstream consumer from enjoying better air. Either all downstream consumers enjoy better air quality or nobody does. We can only "exclude" all consumers together simply by not building the plant in the first place.

We first ask when it is efficient to provide the public good. To answer this question we have to compare the plant's social benefits with its costs. To make this comparison it is useful to have some notation. We denote the decision to built the plant by $q = 1$ whereas $q = 0$ means the retention of the status quo. To build the plant costs 8 units of money. Using the above notation we may denote the plant's costs by $c(q)$ with $c(0) = 0$ and $c(1) = 8$.

There are two goods that consumers enjoy: air quality q and money y. A consumer's preferences are given by $V(q, y) = v(q) + y$; preferences are thus additively separable in air quality and money.[1] We normalize $v(0) = 0$; $v(1) := v > 0$ thus measures the consumers' willingness-to-pay (or their reservation price) for the plant. The consumer would be indifferent between enjoying better air at a cost v to herself and suffering from offensive smells. For most of the paper we assume that there are two downstream consumers. Consumer 1's reservation price is denoted by v_1 and consumer 2's reservation price is v_2. Each consumer's reservation price may take the values 1 or 9 so that four combinations of reservation prices are possible.

The value of the public good to all agents together is the sum of the individual valuations. The condition under which it is efficient to build the plant now follows immediately: If the sum of the consumers' reservation prices exceed the costs, the plant should be built;

otherwise, the status quo should be kept (Samuelson (1954)). At this point it is helpful to think in terms of a decision rule rather than a single decision. By a *decision rule* we understand a function $q(\cdot)$ indicating for each constellation of the relevant parameters (v_1, v_2, c) whether or not to build the plant. The *efficient decision rule* is then given as

$$q(v_1, v_2, c) = \begin{cases} 0, & \text{if } v_1 + v_2 < 8; \\ 1, & \text{otherwise.} \end{cases}$$

Consequently, it is efficient to build the plant if at least one consumer has the high willingness-to-pay; or, to put it differently, the only case in which the status quo should be retained is when both consumers have the low reservation price of 1.

III. The Incentive Problem

The efficient decision rule concerning the provision of a public good depends on a collective valuation — the sum of the consumers' willingness-to-pay. The fact that a social value forms the basis for the efficient provision of the sewage treatment plant creates incentive problems connected with the revelation of consumer preferences. Let us, naturally, assume that an agent's willingness-to-pay is her *private information*. That is, the precise value of v_1 is known only to consumer 1 herself. The planner and consumer 2 only know that consumer 1's willingness-to-pay may be 1 or 9; they do not know which of the two reservation prices applies. Analogously, the precise value of v_2 is consumer 2's private information. In such a situation of *incomplete information* the following question arises: Why should an agent announce her true willingness-to-pay or, in jargon, her *type*?[2]

Suppose, e.g., that the planner simply asks the agents to announce their willingness-to-pay. The plant will be built if the sum of the reported reservation prices exceeds the costs which in turn will be financed by, say, outside sources. Then both consumers announce a reservation price of 9 independent of their true willingness-to-pay. If an agent reports a willingness-to-pay of 1, she runs the risk that the other agent also announces 1. The plant will not be built and the agent loses 1 resp. 9 utils, depending on her true type. The agent rules out this loss by announcing a reservation price of 9. Even if an agent has a valuation of 1, she will announce a willingness-to-pay of 9. Consequently, the above procedure may lead to overreporting. The plant will be built even when it is inefficient to do so, i.e., if $v_1 = v_2 = 1$.

To give the agents an incentive not to overreport, the planner could charge a price for announcing a high reservation price. Consider the arrangement where the planner charges from each consumer, say, half of her announced reservation price. This procedure makes overreporting expensive — unfortunately so expensive that the agents may have an incentive to underreport. Suppose, e.g., consumer 1 believes that consumer 2 announces 9 so that the plant will be built anyway. Then it is optimal for consumer 1 to report a reservation price of 1 (independent of her type) in order to pay as little as possible. Therefore, if agent 1 has a true willingness-to-pay of 9 she will underreport; in jargon we say that consumer 1 is *free riding* on her colleague. If consumer 2 also believes that her colleague announces 9, she will also report a valuation of 1 independent of her true type. The free rider behavior then causes the plant never to be built.

IV. The Revelation Principle

We have just seen that two particular procedures do not implement the efficient decision rule. Do these results indicate that there are no institutional arrangements permitting the efficient provision of public goods? Or can we design institutions, *mechanisms* in economic parlance, which induce consumers to reveal how much the public good is worth to them so that this information can be used to implement the efficient decision rule? The answer to these questions is known as the problem of *mechanism design* for the efficient provision of public goods. In the remainder of this paper, we survey the results achieved by the theory of mechanism design. Some of these results are positive, indicating possibilities for implementing the efficient decision rule; other results are negative, indicating an impossibility of the efficient provision of pure public goods under certain circumstances.

In the theory of mechanism design we look for institutions resulting in the efficient provision of the sewage treatment plant under incomplete information. This looks like a formidable task — one can cook up institutional arrangements however complex. Yet, the so-called *revelation principle* (Dasgupta, Hammond, and Maskin (1979), Myerson (1985)) allows us to pin down the class of mechanisms to search in. Consider the result – the equilibrium outcome in jargon – of some (however complex) mechanism, i.e., whether the plant is built, the agents' tax bills, etc. The revelation principle says that there is another mechanism under which 1) the agents *only* report a willingness-to-pay and under which 2) *truthtelling* is an equilibrium that mimics the equilibrium outcome of the mechanism we started out with. Loosely speaking, the revelation principle tells us that whatever can be done with any mechanism can also be done with a *direct revelation mechanism*, i.e., a mechanism in which the agents only report a reservation price and under which truthtelling is an equilibrium.

Why does the revelation principle hold? Consider the following two scenarios:

Scenario 1: The planner announces a mechanism. Both consumers work out their optimal (possibly highly complex) strategies under the mechanism.

Scenario 2: The planner announces the same mechanism as in Scenario 1. Moreover, she proposes to save the agents the effort of working out what strategies to play under the mechanism. She will do the job for them. The agents only need to tell her their willingness-to-pay and, for each type of consumer she will use the equilibrium strategy that the consumer chooses in Scenario 1.

Obviously, the mechanism in Scenario 2 is a direct mechanism — the agents only report a reservation price. Furthermore, truthtelling is an equilibrium strategy for the consumers. Suppose, e.g., that consumer 2 with a true reservation price of 9 reports a willingness-to-pay of 1 under the first Scenario because this generates the best outcome for her given the strategy of her colleague. If consumer 2 wants the same outcome under the second Scenario, she must tell the planner her true reservation price of 9. The planner will then pick the same strategy as the type 9 consumer 2 does under the first Scenario, i.e., report 1. Consequently, truthtelling is an equilibrium in Scenario 2.

Due to the revelation principle we can confine our attention to the class of direct revelation mechanisms. In particular, the revelation principle is extremely useful to derive negative results: If there is no mechanism with the desired efficiency properties in the class of direct revelation mechanisms, then there is no such mechanism in the class of all mechanisms. Yet before we turn to the impossibility results, we want to describe two positive, i.e., constructive results indicating possibilities for the efficient provision of public goods.

V. The Clarke–Groves Mechanism

Let us start with a famous direct revelation scheme, the so-called Clarke–Groves mechanism (Vickery (1961), Clarke (1971), Groves (1973), Green and Laffont (1979)). The mechanism works as follows: The agents simultaneously announce reservation prices $\hat{v}_1 \in \{1, 9\}$ and $\hat{v}_2 \in \{1, 9\}$. To clarify notation: v_i stands for the true reservation price and $\hat{v}_i$ is the reported valuation; accordingly, if $\hat{v}_i = v_i, i = 1, 2$, the agents tell the truth. The Clarke–Groves mechanism is thus a direct scheme — the agents only announce reservation prices. The sewage treatment plant is provided according to the decision rule

$$q(\hat{v}_1, \hat{v}_2, c) = \begin{cases} 0, & \text{if } \hat{v}_1 + \hat{v}_2 < 8; \\ 1, & \text{otherwise.} \end{cases}$$

Therefore, if the agents tell the truth under the mechanism, i.e., if $\hat{v}_1 = v_1$ and $\hat{v}_2 = v_2$, the plant will be provided efficiently. Finally, conditional on their reported values the consumers pay taxes $t_1(\hat{v}_1, \hat{v}_2)$ and $t_2(\hat{v}_1, \hat{v}_2)$ that are given by Table 1.

It turns out that truthtelling is a *weakly dominant strategy* for both consumers, i.e., irrespective of the other agent's report, telling the truth is never worse than lying. Consider, e.g., consumer 1's payoffs when she has the true willingness-to-pay of 9, see Table 2. If agent

Consumer 2

		$\hat{v}_2 = 1$	$\hat{v}_2 = 9$
Consumer 1	$\hat{v}_1 = 1$	0;0	0;7
	$\hat{v}_1 = 9$	7;0	0;0

Table 1: *Taxes* $(t_1(\hat{v}_1, \hat{v}_2); t_2(\hat{v}_1, \hat{v}_2))$ *under the Clarke–Groves Mechanism.*

Consumer 2

		$\hat{v}_2 = 1$	$\hat{v}_2 = 9$
Consumer 1	$\hat{v}_1 = 1$	0	9
	$\hat{v}_1 = 9$	2	9

Table 2: *Consumer 1's payoffs when she has the true willingness-to-pay* $v_1 = 9$.

2 announces $\hat{v}_2 = 1, \hat{v}_1 = 9$ is strictly better than $\hat{v}_1 = 1$; if agent 2 reports $\hat{v}_2 = 9, \hat{v}_1 = 9$ is no worse than $\hat{v}_1 = 1$. Hence, truthtelling is a weakly dominant strategy; it is straightforward to check that this property holds for all types of consumers. Since both agents tell the truth, under the Clarke–Groves decision rule the plant will be built when it is efficient to do so.

Agent 1 is taxed when she announces 9 and agent 2 reports 1. In this situation the plant would not be built if consumer 1 did not report the high reservation price. By reporting 9 consumer 1 changes the social decision and is, therefore, pivotal. The Clarke–Groves mechanism charges an agent when she is pivotal. In our example there are just two constellations where an agent is pivotal, i.e., when one consumer has the low and the other consumer the high reservation price.

The amount we charge is constructed as follows: By being pivotal, an agent causes the plant to be built which costs society 8. The other consumer gains 1 so that the pivotal consumer imposes a net cost of 7 on society, which is exactly what she has to pay. By charging the consumer the cost she inflicts on society, the consumer internalizes the social consequences of her decision. Under the Clarke–Groves mechanism each consumer's decision problem boils down to the social decision problem and, therefore, the mechanism implements the efficient decision rule.

Unfortunately, the Clarke–Groves mechanism has a major fault. Whenever the plant is built, the agents' payments fail to cover the cost $c(1) = 8$. The mechanism thus needs subsidies to induce agents to tell the truth—money that is no longer available for useful purposes. Since the mechanism needs resources to find out the truth, it is inefficient. Ideally, we would like to have a mechanism where taxes add up to the plant's cost. Nevertheless, it turns out that this is, in general, not possible.[3] Yet it is possible to design a tax scheme where the planner ends up with an expected surplus. Let us briefly sketch how this works. The basic insight is as follows: we can add an extra amount to an agent's tax depending only on what the other agent does without changing the first agent's incentives. To be more specific, the taxes $T_1(\hat{v}_1, \hat{v}_2) = t_1(\hat{v}_1, \hat{v}_2) + h_1(\hat{v}_2)$ and $T_2(\hat{v}_1, \hat{v}_2) = t_2(\hat{v}_1, \hat{v}_2) + h_2(\hat{v}_1)$ give rise to the same incentives as the taxes $t_1(\cdot)$ and $t_2(\cdot)$ do. If we set, say, $h_1(1) = h_2(1) = 0$ and $h_1(9) = h_2(9) = 4$, the tax scheme is given by Table 3.

		Consumer 2	
		$\hat{v}_2 = 1$	$\hat{v}_2 = 9$
Consumer 1	$\hat{v}_1 = 1$	0;0	4;7
	$\hat{v}_1 = 9$	7;4	4;4

Table 3: *Taxes* $(T_1(\hat{v}_1, \hat{v}_2); T_2(\hat{v}_1, \hat{v}_2))$ *under the Clarke–Groves mechanism.*

Now the planner has a balanced budget when both agents have the same reservation price; yet she earns a surplus of 3 when the agents' reservation prices differ. If this surplus is spent on anything affecting the agents' well-being, and if the agents anticipate this, we no longer have the desired incentive structure. We then have to include in the agents' utilities the value they attach to the uses to which the surplus is put. Accordingly, the planner must find a use for the surplus that is of no benefit (or detriment) to the agents — she must, for instance, burn it. Because of this "wasted" tax payment, the allocation of air quality and money will not be efficient.

The next question to ask is, therefore, whether there is an incentive compatible mechanism which is both efficient and budget balanced. This question brings us to another classic scheme, the so-called d'Aspremont–Gérard-Varet mechanism.[4] The answer is yes — if we use a less demanding incentive concept.

VI. The d'Aspremont–Gérard-Varet Mechanism

At this point we have to specify the stochastic properties of the consumers' reservation prices. We assume that each consumer's willingness-to-pay is a random variable $\tilde{v}_1$ and $\tilde{v}_2$, resp., which may both take the values 1 and 9 with probability 1/2. The random variables $\tilde{v}_1$ and $\tilde{v}_2$ are independent.

Under the d'Aspremont–Gérard-Varet mechanism each agent is, essentially, paid the expected value of the other agents' net surpluses conditional on her own report. Then each agent again internalizes the social surplus and has no incentive to distort the decision by manipulating her announcement given that the other agents report the truth. To be more specific, the agents announce reservation prices $\hat{v}_1, \hat{v}_2$ and the public good is provided according to the decision rule

$$q(\hat{v}_1, \hat{v}_2, c) = \begin{cases} 0, & \text{if } \hat{v}_1 + \hat{v}_2 < 8; \\ 1, & \text{otherwise.} \end{cases}$$

Taxes are given by Table 4.[5]

		Consumer 2	
		$\hat{v}_2 = 1$	$\hat{v}_2 = 9$
Consumer 1	$\hat{v}_1 = 1$	0;0	2.5;5.5
	$\hat{v}_1 = 9$	5.5;2.5	4;4

Table 4: *Taxes* $(t_1(\hat{v}_1, \hat{v}_2); t_2(\hat{v}_1, \hat{v}_2))$ *under the d'Aspremont–Gérard-Varet mechanism.*

Now consider, for instance, consumer 1's payoffs when she has a reservation price $v_1 = 9$, see Table 5. First note that truthtelling is not a dominant strategy; if consumer 2 announces $\hat{v}_2 = 9, \hat{v}_1 = 1$ is the best response for the high type consumer 1. Yet, suppose consumer 1 believes that consumer 2 will tell the truth. This means, consumer 1 expects $\hat{v}_2 = 1$ and $\hat{v}_2 = 9$ with equal probability. Then consumer 1's expected payoff from announcing $\hat{v}_1 = 1$ equals 3.25, while her expected payoff from reporting $\hat{v}_1 = 9$ equals 4.25. Thus, if consumer 2 tells the truth, the best reply of the high valuation consumer 1 is also to tell the truth. It is straightforward to check that this self-enforcing property holds for all types of consumers. We may thus conclude that truthtelling is a *Bayes–Nash equilibrium* under the decision rule $q(\cdot)$ and taxes $t_1(\cdot)$ and $t_2(\cdot)$, i.e., if everybody else tells the truth, the best an agent can do is also to tell the truth. Since everybody tells the truth, under the decision rule $q(\cdot)$ the sewage treatment plant will be provided efficiently. Finally note that we have budget balance, i.e., taxes add up to the plant's cost $c(q)$.

The d'Aspremont–Gérard-Varet mechanism uses a less demanding incentive concept than the Clarke–Groves mechanism does. Under the d'Aspremont–Gérard-Varet mechanism truthtelling is optimal for a consumer given that the other agents tell the truth, whereas

Consumer 2

		$\hat{v}_2 = 1$	$\hat{v}_2 = 9$
Consumer 1	$\hat{v}_1 = 1$	0	6.5
	$\hat{v}_1 = 9$	3.5	5

Table 5: *Consumer 1's payoffs when she has the true reservation price* $v_1 = 9$.

under the Clarke–Groves mechanism truthtelling is optimal irrespective of what the other agents do. Under the d'Aspremont–Gérard-Varet mechanism an agent must assess what her fellow agents will do which is not necessary under the Clarke–Groves mechanism. Moreover, to compute her own expected payoffs and thus her optimal strategies an agent must know the probabilities with which her colleague's types are drawn. The agent does not need this bit of information under the Clarke–Groves mechanism. Loosely speaking, under the d'Aspremont–Gérard-Varet mechanism the agents need more information and have more work to do to find their optimal strategies than under the Clarke–Groves mechanism. Technically, this weaker implementation notion means that the class of schemes which are efficient and incentive compatible is larger than under dominant strategy implementation and, therefore, includes budget balance mechanisms.

VII. The Impossibility Theorem

We have just seen that it is possible to achieve efficiency and budget balance by weakening the incentive criterion from dominant to Bayes–Nash strategies. Nevertheless, the d'Aspremont–Gérard-Varet mechanism suffers from a major defect: some types of agents must be forced to participate. Consider, e.g., consumer 2's payoffs when she is of type $v_2 = 1$. Given that her colleague tells the truth, consumer 2's expected payoff from announcing $\hat{v}_2 = 1$ equals -0.75 whereas the message $\hat{v}_2 = 9$ generates an expected payoff of -3.75. Truthtelling is thus better than lying. Recall, however, that consumer 2's payoff in the status quo $v_2(0) = 0$. Consumer 2 with $v_2 = 1$, therefore, expects to do worse under the d'Aspremont–Gérard-Varet mechanism than she does in the status quo. She will not voluntarily participate in the scheme — in jargon, participation is not *individually rational* for a low valuation consumer 2.[6] Consequently, if the agents cannot be forced to participate, the d'Aspremont–Gérard-Varet mechanism will not lead to the efficient provision of the sewage treatment plant.[7]

The next question to ask is, therefore, whether there exist any mechanisms which are efficient, budget balanced, incentive compatible, and individually rational. Unfortunately, the answer to this question is, in general, negative (Myerson and Satterthwaite (1983), Güth and Hellwig (1986)). An efficient mechanism is individually rational for consumer 1 with $v_1 = 1$ if her expected payoff is non-negative. Formally, the mechanism must satisfy

$$1/2 - 1/2\,(t_1(1,1) + t_1(1,9)) \geq 0. \tag{IR1}$$

An efficient mechanism is incentive compatible for consumer 1 with $v_1 = 9$ if the expected payoff of announcing $\hat{v}_1 = 9$ is not less than the expected payoff from reporting $\hat{v}_1 = 1$, or

$$9 - 1/2\,(t_1(9,1) + t_1(9,9)) \geq 9/2 - 1/2\,(t_1(1,1) + t_1(1,9))\,. \tag{IC1}$$

The analogous individual rationality (IR2) and incentive compatibility constraint (IC2) must hold for consumer 2.

For the expected tax revenue from the mechanism we have

$$\begin{aligned} 1/4\,(t_1(1,1) + t_1(1,9) + t_1(9,1) + t_1(9,9)) &\;+ \\ 1/4(t_2(1,1) + t_2(9,1) + t_2(1,9) + t_2(9,9)) &\;\leq \\ 1/4\,(t_1(1,1) + t_1(1,9) + 9 + t_1(1,1) + t_1(1,9)) &\;+ \\ 1/4(t_2(1,1) + t_2(9,1) + 9 + t_2(1,1) + t_2(9,1)) &\;\leq\; 5.5, \end{aligned}$$

where the first inequality follows from (IC1) and (IC2) and the second inequality follows from (IR1) and (IR2). Accordingly, we have derived that the expected revenue from an efficient, incentive compatible, and individually rational mechanism cannot exceed 5.5. Yet, the expected costs given the efficient provision are $3/4 \cdot 8 = 6$. Consequently, in our example any efficient, individually rational revelation mechanism is not budget balanced.

The idea behind the result is the following: An efficient mechanism being individually rational cannot charge too much from the low valuation agents; otherwise, they would not participate in the scheme. Moreover, taxes may not increase too steeply if a consumer reports a high instead of a low reservation price; otherwise, high valuation agents would have no incentive to tell the truth. It turns out that the two constraints on taxes taken together may lead, as in our example, to an insufficient expected revenue, i.e., the mechanism is not budget balanced.

The Myerson–Satterthwaite impossibility theorem tells us that in general no mechanisms exist that are incentive compatible, individually rational, budget balanced, and efficient. Stated differently, if we insist on budget balance, an incentive compatible and individually rational mechanism will typically be inefficient: the public good will be undersupplied. To be more specific, consider a *second-best mechanism* for our example, i.e., a scheme which maximizes the ex ante expected surplus under the constraints of budget balance, individual rationality, and incentive compatibility. A symmetric second-best mechanism is given by the decision rule

$$q(\hat{v}_1, \hat{v}_2, c) = \begin{cases} 0, & \text{if } \hat{v}_1 + \hat{v}_2 = 2; \\ 5/6, & \text{if } \hat{v}_1 + \hat{v}_2 = 10; \\ 1, & \text{if } \hat{v}_1 + \hat{v}_2 = 18. \end{cases}$$

and the taxes given by Table 6.

If one agent has the low and the other agent has the high willingness-to-pay, the plant will not be built for sure (as would be efficient), but only with probability 5/6. The second-

Consumer 2

Consumer 1		$\hat{v}_2 = 1$	$\hat{v}_2 = 9$
	$\hat{v}_1 = 1$	0;0	5/6; 35/6
	$\hat{v}_1 = 9$	35/6;5/6	4;4

Table 6: *Taxes* $(t_1(\hat{v}_1, \hat{v}_2), t_2(\hat{v}_1, \hat{v}_2))$ *under the second-best optimal mechanism.*

best mechanism thus is inefficient; it entails an undersupply of environmental protection. The decision rule is inefficient for the following reasons: By making it less likely that the good will be provided, the planner needs less expected tax revenue to break even. More importantly, the inefficient decision rule penalizes a high valuation agent for reporting a low reservation price. If an agent reports a low willingness-to-pay, the provision is less likely under the second-best mechanism than it is under the first-best decision rule. Since announcing a low valuation becomes less attractive, high valuation agents are willing to pay higher taxes than under the efficient decision rule.

VIII. The Inefficiency Limit Theorems

We have just seen that a budget balanced mechanism does not provide the plant for sure when it is efficient to do so. The mechanism provides inefficiently when the sum of the reservation prices is not much larger than the plant's costs. Accordingly, we might reason as follows: If the number of beneficiaries increases and, therefore, the sum of the reservation prices, the probability that the plant will be built should also increase. Stated differently, the "inefficiency" of a budget balanced mechanism should decrease with the number of consumers.

This reasoning, however, is only partly correct because there is a countervailing effect.

As under the Clarke-Groves mechanism, an agent is willing to pay more than 1 if and only if she is pivotal in the sense that the announcement of the high valuation increases the probability of provision. Yet, the more consumers there are, the lower the probability that any particular agent will be pivotal; the free rider effect increases with the number of agents.

Accordingly, we have two countervailing effects. On the one hand, increasing the number of consumers increases the sum of the reservation prices and, therefore, the expected revenue from a mechanism. On the other hand, the probability of being pivotal decreases with the number of agents and this lowers expected revenue. It can be shown that the combined effect is as follows: The expected revenue from an incentive compatible and individually rational mechanism increases with the square root of the number of agents.[8] That is, the expected revenue is less than proportional to the number of consumers.

To make statements about the surplus of the mechanism, we have to compare the expected revenue with the cost of provision. Suppose, e.g., that the cost of the sewage treatment plant is proportional to the number of consumers. Since the expected revenue is less than proportional to the number of consumers, costs exceed expected revenue for large communities. Consequently, if we insist on budget balance, the plant cannot be provided efficiently. Indeed, under budget balance the probability that the plant will be built at all goes to zero if the number of consumers becomes large. This is in sharp contrast to efficient provision: If the number of consumers becomes large, the efficient probability of provision tends to one. Stated differently, unless there are substantial economies of scale with respect to the number of beneficiaries, the "inefficiency" of a budget balanced, incentive compatible, and individually rational mechanism increases with the number of consumers. In the limit, the plant will not be built at all even though efficiency requires it to be built for sure. (Güth and Hellwig (1986), (1987), Rob (1989), Mailath and Postlewaite (1990)).

IX. Conclusions

We have seen that the theory of mechanism design has been successful in explaining both the current state of the environment and the difficulties encountered by those attempting to implement environmental protection measures. If the latter are provided by voluntary institutional arrangements à la Coase (1960), typically there will be either an undersupply or we fail to achieve budget balance. If the mechanism does not break even, we need outside subsidies which make environmental protection measures more expensive than they are from a technical point of view. With voluntary arrangements we will, therefore, end up with too little environmental protection or, to put it differently, with too much pollution. If we want efficiency, we have to use the d'Aspremont–Gérard-Varet mechanism. Under this institutional arrangement, the efficient decision rule is implemented *and* the mechanism is budget balanced. Nevertheless, the d'Aspremont–Gérard-Varet mechanism is not individually rational, i.e., it relies on collective coercion. The provision of environmental protection measures under incomplete information, therefore, amounts to a tradeoff between efficiency and voluntary participation.

Concerning the tradeoff between efficiency and coercion, a final remark seems in order. Under the d'Aspremont–Gérard-Varet mechanism, some agents must be forced to participate after they learned their reservation prices. Nevertheless, the agents would voluntarily agree on the d'Aspremont–Gérard-Varet mechanism when they decide ex ante, i.e., before they receive their private information. The d'Aspremont–Gérard-Varet mechanism yields ex post efficiency, hence it also maximizes the ex ante expected size of the pie. Ex ante the agents decide under symmetric information. They have an interest to maximize the size of the pie and distribute it among themselves. Accordingly, they will agree on the d'Aspremont–Gérard-Varet mechanism, possibly with ex ante compensatory transfers. Put differently, suppose that at a constitutional stage the individuals decide on rules that must be abided by at a later operational stage. If the agents decide at the constitutional stage under a veil of ignorance, they will unanimously agree on the d'Aspremont–Gérard-Varet mechanism because it maximizes the ex ante expected surplus (Brennan and Buchanan (1985)).

Endnotes

1) If the agents' preference orderings are unrestricted, the Gibbard–Satterthwaite theorem implies the non-existence of mechanisms with the desired properties (Gibbard (1973), Satterthwaite (1975)). Therefore, we restrict our attention to the class of quasi-linear preferences which, essentially, means that there are no income effects.
2) A consumer's type, her private information, is simply her reservation price. Accordingly, consumer 1 can be of type 1 and 9 and so does consumer 2.
3) In our two person, two reservation price example there happens to exist a mechanism implementing in dominant strategies and being budget balanced. The decision rule is as in the Clarke-groves mechanism. Taxes are based on simple cost-sharing, i.e., for $i = 1, 2$,

$$t_i(\hat{v}_1, \hat{v}_2) = \begin{cases} 0, & \text{if } \hat{v}_1 = \hat{v}_2 = 1; \\ 4, & \text{otherwise.} \end{cases}$$

It is straightforward to check that truthtelling is a weakly dominant strategy and, moreover, the mechanism is budget balanced. Nevertheless, this cost-sharing rule no longer works in more complex situations. Adding, e.g., a third reservation price to our example generates the impossibility of implementing in dominant strategies and achieving budget balance.
4) Arrow (1979) derived this mechanism independently.
5) The taxes in Table 4 are constructed as follows. First compute $v_1'(v_2) = v_1(q(v_1, v_2, c)) - c(q(v_1, v_2, c))/2$ and $v_2'(v_1) = v_2(q(v_1, v_2, c)) - c(q(v_1, v_2, c))/2$. v_1' and v_2' thus denote the agents' reservation prices net of half the cost given the efficient provision. Next compute $E_{\tilde{v}_2}(v_2'(v_1))$ and $E_{\tilde{v}_1}(v_1'(v_2))$ where, e.g., the operator $E_{\tilde{v}_2}$ denotes the expectation with respect to the distribution over v_2' that is generated by the prior distribution of $\tilde{v}_2$. Then taxes are $t_1(\hat{v}_1, \hat{v}_2) = -E_{\tilde{v}_2}(v_2'(\hat{v}_1)) + E_{\tilde{v}_1}(v_1'(\hat{v}_2)) + c(q(\hat{v}_1, \hat{v}_2, c))/2$ and $t_2(\hat{v}_1, \hat{v}_2) = -E_{\tilde{v}_1}(v_1'(\hat{v}_2)) + E_{\tilde{v}_2}(v_2'(\hat{v}_1)) + c(q(\hat{v}_1, \hat{v}_2, c))/2$. By construction, $q(\cdot), t_1(\cdot)$, and $t_2(\cdot)$ are efficient, incentive compatible, and budget balanced.
6) The Clarke–Groves mechanism with taxes not yielding any losses also fails to satisfy individual rationality.
7) Emons and Sobel (1991) consider the related problem of designing liability rules. They show that if we use the weaker ex ante efficiency concept (in contrast to ex post efficiency in d'Aspremont–Gérard-Varet), no agent needs to pay more than the maximum damage that can be attributed to her behavior. Consequently, if the agents' utility of not participating is less than their utility of participating and being responsible for all damages associated with their behavior, the mechanism derived by Emons and Sobel is individually rational.
8) The derivation of these limit results is beyond the scope of this paper. The formal arguments behind the theorems are, by and large, those which are used to establish the laws of large numbers.

References

K. J. ARROW, "The Property Rights Doctrine and Demand Revelation under Incomplete Information", in *Economics and Human Welfare: Essays in Honor of Tibor Scitvosky*, (M. Boskin, Ed.), New York: Academic Press (1979), 23 - 39.

C. d'ASPREMONT and L. A. GÉRARD-VARET, "Incentives and Incomplete Information", 11 *Journal of Public Economics* (1979), 25 - 45.

G. BRENNAN and J. BUCHANAN, *The Reason of Rules*, Cambridge: Cambridge University Press (1985).

E. CLARKE, "Multipart Pricing of Public Goods", 11 *Public Choice* (1971), 17 - 33.

R. COASE, "The Problem of Social Cost", 3 *Journal of Law and Economics* (1960), 1 - 44.

P. DASGUPTA, P. HAMMOND, and E. MASKIN, "The Implementation of Social Choice Rules; Some General Results on Incentive Compatibility", 46 *Review of Economic Studies* (1979), 185 - 216.

W. EMONS and J. SOBEL, "On the Effectiveness of Liability Rules when Agents are not Identical", 58 *Review of Economic Studies* (1991), 375 - 390.

A. GIBBARD, "Manipulation for Voting Schemes", 41 *Econometrica* (1973), 587 - 601.

J. GREEN and J. J. LAFFONT, *Incentives in Public Decision Making*, Amsterdam: North Holland (1979).

T. GROVES, "Incentives in Teams", 51 *Econometrica* (1973), 7 - 45.

W. GÜTH and M. HELLWIG, "The Private Supply of a Public Good", 5 *Journal of Economics* (1986), 121 - 159.

W. GÜTH and M. HELLWIG, "Competition versus Monopoly in the Supply of Public Goods", in *Efficiency, Institutions, and Economic Policy*, (R. Pethig and U. Schlieper, Eds.), pp. 183 - 217, Heidelberg: Springer (1987).

G. J. MAILATH and A. POSTLEWAITE, "Asymmetric Information Bargaining Problems with Many Agents", 57 *Review of Economic Studies* (1990), 351 - 367.

R. B. MYERSON, "Bayesian Equilibrium and Incentive-Compatibility: An Introduction", in *Social Goals and Social Organization, Essays in Memory of Elisha A. Pazner*, (L. Hurwicz, D. Schmeidler and H. Sonnenschein, Eds.), pp. 229 - 259, Cambridge: Cambridge University Press (1985).

R. B. MYERSON and M. A. SATTERTHWAITE, "Efficient Mechanisms for Bilateral Trading", 29 *Journal of Economic Theory* (1983), 265 - 281.

R. ROB, "Pollution Claim Settlements under Private Information", 47 *Journal of Economic Theory* (1989), 307 - 333.

P. A. SAMUELSON, "The Pure Theory of Public Expenditure", 37 *Review of Economics and Statistics* (1954), 387 - 389.

M. SATTERTHWAITE, "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions", 10 *Journal of Economic Theory* (1975), 187 - 217.

W. VICKREY, "Conterspeculation, Auctions, and Competitive Sealed Tenders", 16 *Journal of Finance* (1961), 8 - 37.