Collectively Incentive
Compatible Tax Systems

Felix Bierbrauer

Collectively Incentive Compatible Tax Systems

Felix Bierbrauer

September 2006

Collectively Incentive Compatible Tax Systems

Felix Bierbrauer[*]

MPI Collective Goods, Bonn, Germany

This version: September 27, 2006

Abstract

This paper assumes that individuals possess private information both about their abilities and about their valuation of a public good. Individuals can undertake collective actions on order to manipulate the tax system and the decision on public good provision. Consequently, an implementable scheme of taxation has to be collectively incentive compatible. If preferences are additively separable, then an implementable tax systems has the following properties: (i) tax payments do not depend on public goods preferences and (ii) there is no scope for a collective manipulation of public goods preferences. For a quasilinear economy, the optimal tax system is explicitly characterized.

Keywords: Optimal Taxation, Public Good Provision, Revelation of Preferences, Information Aggregation.

JEL: D71, D82, H21, H41

1 Introduction

This paper introduces the concept of a *collectively incentive compatible tax system* as a tool that makes it possible to study two incentive problems simultaneously. The first incentive problem stems from the fact that individuals have private information on their earning ability. This restricts the set of admissible tax systems in a way analyzed in the theory of optimal income taxation, in the tradition of Mirrlees (1971). The second incentive problem arises because individuals have private information on their valuation of a public good. This yields the classical free-rider problem in public good provision and restricts the set of admissible provision rules for public goods.

The joint treatment of these two incentive problems fills a gap in the theory of public economics. This gap exists because the normative theory of public good provision has two separate branches. On the hand, there is the *theory of optimal taxation.* This theory assumes that there is a large economy and that the tax setting institution can be assumed to know the distribution of characteristics in the economy. It solves for an optimal scheme of taxation and public good provision under a public sector budget constraint. The optimal quantity of a public good is determined according to a modified version of the *Samuelson rule*, after Samuelson (1954), which takes the use of distortionary tax instruments to finance public good provision into account.[1]

In this approach there is no problem of information aggregation. The economy is large. This justifies the assumption, that the distribution of characteristics is taken to be commonly known. Consequently, there is no need to elicit individual valuations of public goods.

The second branch of the literature on public goods provision is driven by this latter problem. I refer to it as the *mechanism design* approach.[2] In the simplest setting, a benevolent mechanism designer has to choose a provision rule for a public good and a payment scheme. An optimal decision on public good provision requires that information on the distribution of preferences is made available, i.e there is a problem information aggregation.

Such an analysis is typically undertaken for an economy consisting of finitely many individuals. Under an optimal provision rule, each single individual can affect the decision on public good provision and hence the enjoyment of the public good by all other individuals. Public good provision thus becomes

---

[1] Examples of this approach include Atkinson and Stern (1974), Wilson (1991), Boadway and Keen (1993), Nava et al. (1996), Sandmo (1998), Hellwig (2005, 2004) and Gaube (2000, 2005).

[2] This literature originates from the study of Vickrey-Clarke-Groves Mechanisms, see Clarke (1971) and Groves (1973). A survey can be found in Laffont (1987) or the textbook of Mas-Colell et al. (1995). Recent contributions to this line of research are Hellwig (2003) and Norman (2004).

a rather complex strategic game driven by multilateral externalities.

To summarize this brief overview, the idea that a reasonable criterion for public good provision requires to collect information on individual valuations of the public good and that this causes an incentive problem, has been addressed in finite economies but not in large economies and not in conjunction with the tax instruments which are used to finance public expenditures. The present paper provides a framework that addresses these issues.

At a conceptual level, this raises the question of an appropriate solution concept. The main issue is whether information aggregation is really an incentive problem in a large economy. To see this, suppose there are infinitely many individuals, each with private information on the own valuation of the public good. Moreover, assume that the institution that decides on public good provision uses a revelation game to learn about the distribution of preferences. As the economy is large, no single individual has a direct impact on the distribution of preferences. Hence, no individual has a reason to hide his valuation of the public good, implying that information aggregation does not involve an incentive problem.

The present paper, however, takes a different view. It is assumed that individuals can form coalitions in order to manipulate the decision on public good provision. Consequently, information on the distribution of public goods preferences can be acquired only if there is no large group of individuals who gain from a collective manipulation of announced public goods preferences. Allocations which do not provoke such strategic manipulations by groups of agents are henceforth called *collectively incentive compatible*.

The formal model is based on the environment analyzed in the theory of optimal income taxation, following Mirrlees (1971). The new assumptions introduced in this paper are *first*, that individuals do not only have private information about their ability level but also about their valuation of a public good and *second*, that there is *aggregate uncertainty* because the joint cross-section distribution of earning abilities and public goods preferences is not commonly known.

The characterization of the set of implementable allocation is treated as a problem of mechanism design in a large economy.[3] An allocation consists of an income tax schedule and a provision rule for public goods. To be implementable it has to fulfill three requirements. First, it has to be *budgetary feasible*. Second, it has to be *individually incentive compatible (I-IC)*: From a single individual's perspective there is no reason to hide the own characteristics – taking the announcements of all other individuals in the revelation game as given. Finally, it has to be *collectively incentive compatible (C-IC)*:

---

[3]This approach has been introduced by Hammond (1979) and Guesnerie (1995). See Hellwig (2004) for a recent contribution.

2

No coalition of individuals has an incentive to manipulate the profile of announcements in order to affect the decision on public good provision, taking the announcements of individuals outside the coalition as given.

The main result of the paper provides a characterization of individually and collectively incentive compatible tax systems under the assumption that individuals' utility functions are additively separable between private and public goods. *First*, the tax system can not use differences in public goods preferences as a screening device. Put differently, individual tax payments are only a function of ability levels and do not depend on an individual's taste for public goods. *Second*, the individual and the collective dimension of the incentive problem are also separable in the following sense. *C-IC* holds if no coalition of individuals gains from a manipulation of public goods preferences, taking as given that these individuals reveal their earning ability. Put differently, there is no need to worry about coalitions that manipulate the announced profile of earning ability. A revelation of individual ability levels is in turn ensured by the requirement of *I-IC*, taking as given that that there is no collective manipulation of public goods preferences.

This observation proves convenient for a more explicit characterization of implementable allocations in more specific environments.[4] To illustrate this, one such application is studied in more detail, namely an economy in which individuals have quasilinear preferences over the quantity of a public good and their individual payment obligation.[5]

The remainder of the paper is organized as follows. Section 2 contains the formal description of the economy. In addition, the example of a quasilinear economy is used to demonstrate that an optimal scheme of income taxation and public good provision is in general vulnerable to the formation of manipulating coalitions. In section 3 the solution concept of a *collectively incentive compatible tax system* is introduced. This section also contains a discussion of the related literature on mechanism design problems under coalition formation. Section 4 derives the result that, with separable preferences, a separation of *individual* and *collective* incentive problems is possible. In Section 5 this observation is used to characterize the optimal *I-IC* and *C-IC* allocation in the quasilinear economy. The last section contains concluding remarks. All proofs are in the appendix.

---

[4]See the applications in Bierbrauer (2006) and Bierbrauer (2005).

[5]For this environment, Bierbrauer and Sahm (2006) characterize the optimal provision rule for a public goods under the premise that information on the distribution of preferences is made available by a voting procedure.

2 The Environment

There is a large set of individuals identified with the unit interval $I = [0,1]$ and equipped with measure μ. An individual $i \in I$ has a utility function U defined over the quantity $Q \in \mathbb{R}_+$ of a non-excludable public good, and bundles of private goods $A \in \mathbb{R}^l$. In addition, utility depends on individual characteristics. I distinguish a taste parameter $\theta_i \in \Theta$, $\Theta \subset \mathbb{R}_+$, to formalize heterogeneity regarding valuations of the public good and a productivity or skill parameter $w_i \in W$, $W \subset \mathbb{R}_+$. For brevity, I denote a pair of individual characteristics (θ_i, w_i) by γ_i and the set $\Theta \times W$ by Γ. U is thus written as

$$U = U(Q, A, \gamma_i) .$$

Example 1 In the theory of optimal income taxation A is a pair (C, Y) consisting of consumption of private goods $C \in \mathbb{R}_+$ and effective labour supply or income $Y \in \mathbb{R}_+$. In this setting, the productivity parameter captures individual heterogeneity with respect to the utility loss associated with a given level of effective labour supply.

When discussing applications, I impose the assumption that the utility function U is additively separable in the utility contribution of the public good, depending on the taste parameter θ_i, and the utility contribution of A, depending on the skill parameter w_i.

Assumption 1 The utility function U is additively separable:

$$U = v(Q, \theta_i) + u(A, w_i) .$$

The assignment of characteristics to individuals is represented by an assignment function $\gamma_a : I \to \Gamma$ with image denoted by $\{\gamma_i\}_{i \in I} = \{(\theta_i, w_i)\}_{i \in I}$. It is assumed throughout that there is *assignment uncertainty*; i.e. the function γ_a – or equivalently the profile $\{\gamma_i\}_{i \in I}$ – is not commonly known. Instead, individual i has private information on the parameter γ_i.

Assumption 2 Almost all assignments γ_a are measurable functions.

Assumption 2 implies that expressions such as e.g.

$$\mu(\{i \mid \theta_i \leq \theta \text{ and } w_i \leq w\}) \quad \text{or} \quad \mu(\{i \mid \theta_i \leq \theta \})$$

are, for any resolution of assignment uncertainty γ_a, well defined.
In addition to assignment uncertainty, there is *aggregate uncertainty* referring to the empirical distribution of individual characteristics in the economy. From an ex ante perspective there are different states of the economy.

4

Each such state corresponds to a cross-section distribution of characteristics and is represented by a cumulative distribution function (cdf) $D : \Gamma \to [0, 1]$ that lists for each $\gamma = (\theta, w)$ the fraction of individuals with characteristics $\gamma_i \leq \gamma$,

$$D(\gamma) = \mu(\{i \mid \theta_i \leq \theta \text{ and } w_i \leq w\}) .$$

Assumption 3 There is aggregate uncertainty, in the sense that the actual cross-section distribution of characteristics D in the economy is not commonly known. There is a commonly known set $\mathcal{D}$ of possible states of the economy.

The following information about the distribution of characteristics in the economy is common knowledge. There is *aggregate stability* regarding the marginal distribution of productivity parameters.

Assumption 4 There is aggregate stability with respect to the productivity parameter; i.e. for any $D \in \mathcal{D}$ there is a commonly known marginal cumulative distribution function $F : W \to [0, 1]$ with $F(w) = \mu(\{i \mid w_i \leq w\})$.

Remark 1 At this general level, there is no need to be more specific on the relation between randomness at the individual level – i.e. the precise nature of assignment uncertainty – and Assumptions 3 and 4 on the aggregate structure of the economy. In the literature one often finds that $\{\gamma_i\}_{i \in I}$ is taken to be the realization of stochastic process consisting of independent and identically distributed *(i.i.d.)* random variables. In addition, with appeal to some *Law of Large Numbers for Large Economies*, any realization of this process is assumed to induce an assignment that is consistent with the assumptions imposed on the aggregate features of the economy. A mathematical foundation for this approach is provided by Al-Najjar (2004).

The following example is used repeatedly to illustrate the main ideas of this paper.

2.1 An Example

Let $W = [\underline{w}, \bar{w}]$ be a compact interval and let F be such that there exists a density f that is strictly positive on $(\underline{w}, \bar{w})$. Let $\Theta = \{\theta_L, \theta_H\}$ with $\theta_L < \theta_H$. *Aggregate uncertainty* is formalized as follows: Denote by p the fraction of individuals with a high taste parameter, $p = \mu(\{i \mid \theta_i = \theta_H\})$. While each individual observes the own taste realization, p is an unknown parameter between 0 and 1. It is assumed that p is the only source of aggregate

uncertainty; that is, there exists a bijection between possible values of the parameter $p \in [0,1]$ and the set $\mathcal{D}$ of states of the economy. To be more precise, the following assumptions are imposed.

Assumption 5 The assignment of characteristics to individuals proceeds sequentially. First, there is a skill assignment $w_a : W \to I$. Second, there is an assignment of taste parameters to skill parameters $\theta_a : W \to \Theta$ with image denoted by $\{\theta_w\}_{w \in W}$. The interpretation is that for an individual with productivity level w the taste parameter is given by θ_w. Skill and taste assignments are assumed to satisfy the following properties.

i) For any $p \in [0,1]$, the profile $\{\theta_w\}_{w \in W}$ is the realization of an *i.i.d.* process of random variables $\{\tilde{\theta}_w\}_{w \in W}$.

ii) A Law of Large Numbers applies: almost all realizations of $\{\tilde{\theta}_w\}_{w \in W}$ are such that, for every subinterval $[w_1, w_2] \subset W$,

$$\frac{1}{F(w_2) - F(w_1)} \int_{w_1}^{w_2} \theta_w \, dF \quad = \quad p\theta_H + (1-p)\theta_L \; .$$

Remark 2 In Assumption 5 the random taste assignment operates on the set W of possible skill levels and not directly on the set of individuals I. As both W and I are continua this does not affect the mathematical structure. The results of Al-Najjar (2004) remain applicable.

This specification has the property that in every state of the economy the empirical marginal cross-section distribution of the skill parameter and the empirical marginal cross-section distribution of the taste parameter are independent. Put differently, in every state of the economy, the average taste level is the same on every subinterval of W.[6]

2.2 The Taxation Principle

A tax system is interpreted as the outcome of a mechanism design problem under the restriction that allocations have to be anonymous. An *anonymous allocation* consists of two mappings, a provision rule for the public good,

$$Q : \mathcal{D} \to \mathbb{R}_+, D \mapsto Q(D),$$

and a menu of private goods bundles

$$A : \mathcal{D} \times \Gamma \to \mathbb{R}^l, (D, \gamma) \mapsto A(D, \gamma) \; .$$

---

[6] A setup that does not use this assumption is the *Two-Class Economy* analyzed in Bierbrauer (2005).

Remark 3 Following Guesnerie (1995) two aspects of anonymity can be distinguished. There is *recipient anonymity* as the private goods bundle dedicated to an individual depends only on that individual's characteristics but not on the index i. In addition, there is *anonymity in influence*. Neither the menu $\{A(D, \gamma)\}_{\gamma \in \Gamma}$ nor the provision level $Q(D)$ change in response to a permutation of $\{\gamma_i\}_{i \in I}$ that leaves the cross section distribution D unaffected.[7]

Definition 1 An anonymous allocation is said to be *individually incentive compatible (I-IC)* if

$$\forall D, \forall \hat{\gamma}, \forall \gamma : \quad U(Q(D), A(D, \gamma), \gamma) \geq U(Q(D), A(D, \hat{\gamma}), \gamma) .$$

It is *feasible* if for all $D \in \mathcal{D}$, the collection $[Q(D), \{A(D, \gamma)\}_{\gamma \in \Gamma}]$ belongs to the set of feasible allocations Z.

The individual incentive compatibility conditions are stated for a given D; that is, they restrict the possibility for a differential treatment of individuals only within a given cross-section distribution of characteristics. They do not place constraints on the ability of an allocation to specify different outcomes for different members of $\mathcal{D}$. This is due to the fact that in a continuum economy any one individual has a mass of zero and hence does not affect the distribution of characteristics; i.e. there is no impact on the state of the world as perceived by the mechanism designer. In combination with the postulate of anonymity this implies in particular, that no single individual has an impact on public good provision.

Remark 4 It is possible to prove a revelation principle for anonymous allocations. Accordingly, the set of anonymous allocations which are implementable as the outcome of an *anonymous game*[8] in which each individual has a dominant strategy coincides with the set of *I-IC* allocations. A more precise statement and a proof of this revelation principle can be found in the Appendix.

If one seeks for allocations that can be reached via some anonymous game form, then Remark 4 allows to restrict attention to anonymous allocations

---

[7]Guesnerie (1995) argues that the consideration of anonymous allocations contains no loss of generality if the profile of characteristics $\{\gamma_i\}_{i \in I}$ is viewed as the realization of an *i.i.d.* process of random variables. Under this assumption there is no correlation among individual characteristics that a mechanism designer could potentially exploit.

[8]An anonymous game is defined by the property that a player's payoff depends on the own action and the own characteristics, while the actions chosen by other players only enter via their empirical distribution. More details can be found in Kalai (2004).

which are *I-IC* and feasible. Moreover, as the following proposition claims those allocations have the property of being *decentralizable*.

Definition 2 A feasible anonymous allocation $[Q, A]$ is called *decentralizable* if there exists a collection of budget sets $\{B(D)\}_{D \in \mathcal{D}}$ such that

$$\forall \gamma, \forall D : \quad A(D, \gamma) \in \quad \text{argmax}_{X \in B(D)} \quad U(Q(D), X, \gamma)$$

Proposition 1 (Taxation Principle) An anonymous allocation is *I-IC* and feasible if and only if it is a decentralizable.

A proof can be found in Hammond (1979). According to the *taxation principle*, any *I-IC* and feasible allocation has the property of being decentralizable via a budget set $B(D)$ that is common for all individuals in the economy, and vice versa. Consequently, the set of decentralizable allocations is the relevant object for a study of tax systems. Any tax system generates a decentralized allocation, where the budget set $B(D)$ is shaped by the available tax instruments. The final allocation then results from the solution of the utility maximization problems that individuals face under the given tax system. In reverse direction, the *taxation principle* implies that to each *I-IC* allocation one can find a corresponding tax system – implicitly defined as the set of tax instruments that generate the set $B(D)$.

In the theory of optimal income taxation the *taxation principle* takes a more concise form as illustrated by the following example.

Example 2 Suppose that $A = (C, Y)$ as in Example 1. An anonymous allocation $[Q, C, Y]$ is said to be *feasible* if

$$\forall D : \quad \int_{\Gamma} Y(D, \gamma) - C(D, \gamma) \, dD \quad = \quad K(Q(D)) \, ,$$

where $K(\cdot)$ is a cost function that captures the resource requirement of public good provision. The allocation $[Q, C, Y]$ is said to be *decentralizable by an income tax*, if there exists a function $T : \mathcal{D} \times \mathbb{R}_+ \to \mathbb{R}$ such that for all D and for all γ: Consumption equals after tax income: $C(D, \gamma) = Y(D, \gamma) - T(D, Y(D, \gamma))$, individuals choose a utility maximizing level of income subject to the given income tax schedule,

$$Y(D, \gamma) \quad \in \quad \text{argmax}_Y \quad U(Q(D), Y - T(D, Y), Y, \gamma) \, ,$$

and the public sector budget constraint is satisfied,

$$\int_{\Gamma} T(D, Y(D, \gamma)) dD \quad = \quad K(Q(D)) \, .$$

For this environment the *taxation principle* takes the following form: An anonymous allocation is *I-IC* and feasible if and only if it is *decentralizable by an income tax*. A proof can be found in the Appendix.

2.3   Taxes must not depend on public goods preferences

If individuals' utility functions are additively separable, then the tax system cannot exploit heterogeneity with respect to individuals' valuations of the public good.

Proposition 2 Suppose Assumption 1 holds. An anonymous allocation is *I-IC* if and only if it satisfies the following properties:

i) The *no discrimination of taste (NDT)* property:

$$\forall D, \forall w, \forall \theta, \forall \theta' : \quad u(A(D, \theta, w), w) = u(A(D, \theta', w), w) \ .$$

ii) The *individual revelation of productivity (I-RP)* property:

$$\forall D, \forall \theta, \forall w, \forall w' : \quad u(A(D, \theta, w), w) \geq u(A(D, \theta, w'), w) \ .$$

The *I-RP* property requires that individuals are willing to reveal their earning ability, taking a truthful announcement of their taste parameter as given. This is the notion of incentive compatibility that is underlying the literature on optimal income taxation.

The *NDT* property says that two individuals who have the same ability level but differ in their taste parameter must receive the same utility from private goods consumption under an *I-IC* allocation. Put differently, tax payments of individuals must not depend on their views on the desirability of public good provision.

Formally, the *NDT* requirement follows from the fact that individuals take D and hence the level of public good provision as given. Due to the separability assumption, the utility contribution of the public good does not enter individual incentive compatibility constraints; i.e the *I-IC* conditions become independent of taste parameters. Consequently, an admissible tax system can use only individual differences in productivity levels as a screening device and must leave all individuals indifferent with respect to possible taste announcements.

In the following, I will argue that this creates a scope for collective manipulations of taste parameters. Since an *I-IC* allocation can not undermine the willingness of individuals to announce false taste parameters, there is room for collective manipulations in the following sense: a subset of agents with positive mass can affect the announced distribution of public goods preferences – even without violating individual incentive constraints. Again, this is illustrated by means of an example in the next subsection.

2.4 The Example continued

Reconsider the economy described in subsection 2.1. Assume that the utility function satisfies Assumption 1 and, moreover, takes the following quasi-linear form

$$U \ = \ \theta Q - \frac{t}{w} \ ,$$

where t is the individual's contribution to the cost of public good provision. This utility specification captures the idea that less able individuals suffer from a larger utility loss if forced to generate the income that is needed to meet a given payment obligation t.

U is the cardinal utility function that is relevant for welfare assessments. Its ordinal properties are equivalently represented by the following monotone transformation $V := wU = \theta wQ - t$. I refer to to the term θw as the *effective valuation* of the public good by an individual with characteristics (θ, w).

Recall the information structure specified in subsection 2.1. Each state of the economy corresponds to a value for the parameter p that determines the average valuation of the public good $p\theta_H + (1-p)\theta_L$. An anonymous allocation is hence represented by a provision rule for the public good $Q : p \mapsto Q(p)$ and a payment scheme $t : (p, \theta, w) \mapsto t(p, \theta, w)$ that specifies for each state of the economy the contribution of an individual with characteristics (θ, w) to the cost of public good provision.

A straightforward application of Proposition 2 yields the observation that an allocation $[Q, t]$ is *I-IC* if and only if, for each p all individuals have the same payment obligation; i.e. for all p and for all (θ, w) and all (θ', w'), $t(p, \theta, w) = t(p, \theta', w')$. Moreover, assuming that the cost of public good provision is given by a strictly increasing and strictly convex per capita cost function $K : Q \mapsto K(Q)$ and adding a resource constraint yields the result that $[Q, t]$ is *I-IC* and *feasible* if and only if the payment scheme t prescribes *equal cost sharing*; that is, for all p and for all (θ, w), $t(p, \theta, w) = K(Q(p))$. These observations allow to represent an individual's assessment of an allocation rule $[Q, t]$, which is budgetary feasible and incentive compatible, in the following reduced form,

$$V(p, \theta, w) \ := \ \theta wQ(p) - K(Q(p)) \ . \tag{1}$$

In what follows, I consider the choice of an optimal *I-IC* and *feasible* allocation by a benevolent utilitarian planner. The planner evaluates an allocation from the ex ante perspective, i.e. before the actual value of p is known. For simplicity, I assume that the planner takes the actual state of the economy p to be the realization of a random variable $\tilde{p}$ that is uniformly distributed on the unit interval $[0, 1]$.[9] Using the Law of Large Numbers in Assumption

---

[9]Throughout I do not impose a common prior assumption. Only the prior beliefs of the mechanism designer are specified.

5, this implies that expected utilitarian welfare from the ex ante perspective is given by

$$EW \quad := \quad \int_0^1 \left\{ \left[p\theta_H + (1-p)\theta_L \right] Q(p) - \left[\int_{\underline{w}}^{\bar{w}} \tfrac{f(w)}{w} \, dw \right] K(Q(p)) \right\} dp$$

$$= \quad \lambda \int_0^1 \left\{ \bar{v}(p) Q(p) - K(Q(p)) \right\} dp \, ,$$

where $\lambda := \int (1/w) f(w) dw$ is an index of the marginal welfare effect of the cost of public good provision under equal cost sharing and

$$\bar{v}(p) \quad := \quad \frac{p\theta_H + (1-p)\theta_L}{\lambda}$$

is the *effective utilitarian valuation* of the public good.

I will now show that if EW is maximized under the requirements of *I-IC* and feasibility only, then the resulting allocation is vulnerable to manipulative collective actions by groups of individuals who oppose the decision on public good provision. To see this consider the provision rule $Q^* : p \mapsto Q^*(p)$ that is chosen by a utilitarian planner who maximizes EW pointwise; i.e. who maximizes the expression $\bar{v}(p) Q(p) - K(Q(p))$ for every $p \in [0,1]$. This provision rule is characterized by a continuum of first order conditions

$$\forall p : \quad \bar{v}(p) \quad = \quad K'(Q^*(p)) \, .$$

Under Q^* individual preferences over the *"announced state of the world"* can be represented by the following indirect utility function,

$$V^*(p, \theta, w) \quad := \quad \theta w Q^*(p) - K(Q^*(p)) \, .$$

It is easily verified that

$$V_p^*(p, \theta, w) \quad = \quad Q^{*\prime}(p) \Big(\theta w - \bar{v}(p) \Big) \quad \begin{cases} < 0 & \text{if} \quad \theta w < \bar{v}(p) \, , \\ = 0 & \text{if} \quad \theta w = \bar{v}(p) \, , \\ > 0 & \text{if} \quad \theta w > \bar{v}(p) \, . \end{cases}$$

That is, under provision rule Q^* an individual prefers a larger level of p – or equivalently a larger level of public good provision – if and only if the own *effective valuation* exceeds the *effective utilitarian valuation*. Likewise an individual with an *effective valuation* below the average prefers to have a lower quantity of the public good.

These observations imply that groups of individuals would refuse to reveal their true taste parameters if they could thereby affect the mechanism designer's perception of p. To see this, consider the set of individuals with a low taste parameter and a high skill level who have an effective valuation close to $\theta_L \bar{w}$. Moreover, for the sake of concreteness, assume that these

individuals share the belief that p is very low.[10] Put differently, these individuals believe that a vast majority has a low taste realization and that, as a consequence, their own effective valuations lie above the effective utilitarian valuation. Hence, under Q^*, this group of individuals expects that the quantity of the public good is too low and would be happy if the mechanism designer had a larger perception of p. But this implies that these individuals are better off if they collectively announce a high taste parameter.

These considerations highlight the following issues: *First*, a collective deviation from the truth may be beneficial for a subset of agents. *Second* such a collective deviation is not prevented by individual incentive compatibility. Given that all high skilled individuals lie about their taste parameter, there is no incentive for an isolated high skilled individual to reveal his taste parameter truthfully. Due to the *NDT* property, this is a systematic feature. With separable preferences a collective deviation involving taste parameters is not undermined by individual incentives.[11]

3 Collective Incentive Compatibility

As the discussion in the preceding subsection has shown, the requirement of *I-IC* is not sufficient to ensure that an allocation is able to fulfill the task of information aggregation. Under *I-IC*, incentives for a collective manipulation of the mechanism designer's perception of the distribution D are not eliminated. In the following the notion of a *collectively incentive compatible (C-IC)* tax system is introduced that does not suffer from this problem.

More specifically, the definition of *collective incentive compatibility* that is given below requires that truth-telling is a dominant strategy for each conceivable coalition. The main advantage of this approach is that the analysis of coalition formation does not require assumptions on the prior beliefs of individuals. Moreover, as will be explained below, the focus on dominant strategies implies that coalition formation can be analyzed as if individuals had complete information on the state of the economy.

Before the definition of a *C-IC tax system* can be stated, I need to define a *coalition* and a *subcoalition*. For reasons that will become clear, I require that any potentially manipulating subset of agents must have a fixed minimal size $\epsilon > 0$, where ϵ can be arbitrary small. Moreover, a subcoalition J' is a subset of a given coalition J that excludes at least an ϵ - mass of individuals from J.

---

[10]This means that ex interim these individuals have prior beliefs that put a lot of probability mass on values of p which are close to zero.

[11]A further example for the vulnerability of an optimal *I-IC* and *feasible* allocation is found in Bierbrauer (2005).

Definition 3 A *coalition* J is a subset of agents with $\mu(J) \geq \epsilon$, for some fixed but arbitrary $\epsilon > 0$. A *subcoalition* J' of coalition J is a coalition with the properties $J' \subset J$ and $\mu(J') \leq \mu(J) - \epsilon$.

Two implications of this definition, that are used below, are the following:

i) A coalition J with $\epsilon \leq \mu(J) < 2\epsilon$ possesses no subcoalition.

ii) Consider a chain $\ldots \subset J''' \subset J'' \subset J' \subset J$ resulting from a successive formation of subcoalitions. Any such chain has a finite length.

The following notation is needed to describe the potential impact of a coalition on the perceived distribution of characteristics. Denote by $\mathcal{J}_\epsilon$, with typical element J, the set of subsets of I which satisfy $\mu(J) \geq \epsilon$. Denote the true profile of characteristics in J by $\gamma_J := \{\gamma_j\}_{j \in J}$. Denote the reported profile by $\hat{\gamma}_J := \{\hat{\gamma}_j\}_{j \in J}$. Let the actual distribution of characteristics in the economy be $D \in \mathcal{D}$. Denote the cross section distribution of announcements induced by $\hat{\gamma}_J$ if all individuals not in J report truthfully by $\hat{D}(\hat{\gamma}_J, D) \in \Delta_\Gamma$, where Δ_Γ is the set of cdfs with domain Γ.[12]

Consider a coalition J with $\mu(J) \geq 2\epsilon$. Suppose that J induces state perception $\hat{D}(\hat{\gamma}_J, D)$ via the profile of announcements $\hat{\gamma}_J$. Suppose that the members of a subcoalition J' of J deviate from this profile and report instead according to $\tilde{\gamma}_{J'} \neq \hat{\gamma}_{J'}$. The induced announced distribution of characteristics is denoted by $\hat{D}(\tilde{\gamma}_{J'}, \hat{\gamma}_{J \setminus J'}, D)$.

Definition 4 A coalition J is said to *manipulate* an allocation if there exists $D \in \mathcal{D}$, and $\hat{\gamma}_J \neq \gamma_J$ with the following properties:

i) *Undetectability.* The induced distribution is feasible: $\hat{D}(\hat{\gamma}_J, D) \in \mathcal{D}$.

ii) *Unanimity.* All coalition members are strictly better off when choosing to report according to $\hat{\gamma}_J$ instead of γ_J. $\forall j \in J$:

$$U(Q(\hat{D}), A(\hat{D}, \hat{\gamma}_j), \gamma_j) \quad > \quad U(Q(D), A(D, \gamma_j), \gamma_j) .$$

iii) *Individual Stability.* No coalition member departs – unilaterally – from coalitional behavior. Given the *I-IC*-constraints, this requires $\forall j \in J$:

$$U(Q(\hat{D}), A(\hat{D}, \hat{\gamma}_j), \gamma_j) \quad = \quad U(Q(\hat{D}), A(\hat{D}, \gamma_j), \gamma_j) .$$

iv) *Collective Stability.* There does not exist a subcoalition $J' \subset J$, with an *undetectable* collective deviation $\tilde{\gamma}_{J'} \neq \hat{\gamma}_{J'}$ that induces a state perception $\hat{D}(\tilde{\gamma}_{J'}, \hat{\gamma}_{J \setminus J'}, D)$ that makes all members of J' strictly better

---

[12]Assumptions 3 and 4 imply that $\mathcal{D} \subset \Delta_\Gamma$ and $\mathcal{D} \neq \Delta_\Gamma$.

off relative to $\hat{D}(\hat{\gamma}_J, s)$ (*unanimity*), prescribes for all its members individually best responses given the state perception $\hat{D}(\tilde{\gamma}_{J'}, \hat{\gamma}_{J \setminus J'}, D)$ (*individual stability*) and is not threatened by further collective manipulations, which satisfy all these requirements (*collective stability*).

An allocation is said to be *collectively incentive compatible (C-IC)* if there exists no manipulating coalition.

According to this definition, a coalition considers a collective deviation in response to truth-telling of all other individuals. The scope for manipulation is limited by the requirement that it must not be *detectable*; i.e. relevant coalitional plans need to have the property that it does not become apparent that a manipulation has occurred. Moreover, coalition members have to agree unanimously on a deviation and may not use side payments to reach such an agreement. Finally, a coalition has to meet two *stability* requirements. The incentives coalition members face individually must not conflict with the message profile used by the coalition; that is, collective manipulations are a concern only in so far as they do not conflict with *I-IC*. In addition, a conceivable collective manipulation has to be such that it does not provoke the formation of a subcoalition which departs from the original coalitional plan.

A peculiarity of Definition 4 is that collective stability of a coalition J is defined with reference to the collective stability of a subcoalition $J' \subset J$. The requirement of a minimal size for coalitions and subcoalitions ensures that these notions can be traced back to the collective stability of a set of "smallest" coalitions, those with mass between ϵ and 2ϵ.[13]

The requirement of collective incentive compatibility ensures that the allocation $[Q, A]$ can be implemented as the outcome of an anonymous revelation game in such a way that for each coalition J truth-telling is a dominant strategy in the following sense: for any profile of announcements of individuals not in J, truth-telling is the best stable collective announcement for individuals in J.[14] In Remark 4 it has been claimed that *I-IC* of an allocation is equivalent to the possibility to implement it as the outcome of an anonymous revelation game in which each individual possesses a dominant strategy. The requirement of *C-IC* is hence commensurate to *I-IC* in the sense that both ensure implementability in dominant strategies.

---

[13]Bernheim et al. (1986) introduce the notion of a *coalition-proof* Nash-equilibrium for games with a finite number of players. They provide a recursive definition based on a definition of coalition-proofness for games with only one player. The above definitions of *stability* are an adaption of this idea for the present setup.

[14]Alternatively, *C-IC* can be framed as a *robustness*-requirement that ensures incentive compatibility of collective actions irrespective of the prior beliefs of individuals in the economy; see Bergemann and Morris (2005) or Kalai (2004).

The interpretation of these requirements in terms of admissible tax systems is the following. According to the *taxation principle* in Proposition 1, the requirements of anonymity and *I-IC* are equivalent to the existence of a tax system that can be used to decentralize an allocation. However, decentralization via the budget set $B(D)$ presumes that the actual state D has already been determined. The additional requirement of *C-IC* ensures that this information is indeed available; that is, under *C-IC* the tax system does not rely on information that creates a scope for collective manipulations by groups of individuals. Put differently, allocations that are *I-IC* and *C-IC* imply the existence of a tax system and simultaneously allow for information aggregation.

In the remainder of this section I will explain two aspects of the *C-IC* requirement in more detail. The last subsection contains a discussion of related concepts of coalition-proofness that can be found in the literature.

3.1 Off-the equilibrium tax systems

One might take the view that the requirement of *C-IC* is too strong in the sense that there exist alternative ways of achieving a non-manipulable allocation. For instance, a mechanism designer could use use *"off-the-equilibrium rewards"* for subcoalitions to destabilize potential coalitions. To illustrate this, suppose that in state $D \in \mathcal{D}$, coalition J would want to induce state perception $\hat{D} \in \mathcal{D}$ using the false announcements in $\hat{\gamma}_J$. Now suppose that the mechanism designer rewards a further deviation of a subcoalition $J' \subset J$ to some announced distribution $\tilde{D}$, where $\tilde{D}$ does not belong to the set of feasible states $\mathcal{D}$. Thereby the initial manipulation of coalition J is undermined.[15] Moreover it is undermined in a way that is not costly in terms of the welfare properties of the final allocation because the outcome promised to individuals in J' under $\tilde{D}$ is not part of an equilibrium allocation. Hence, an implicit assumption underlying the requirement of *C-IC* is that such *"off-the-equilibrium tax systems"* that only serve to destroy collective manipulations of *"equilibrium tax systems"* can not be used. While this entails a loss of generality, it still seems to be a reasonable way of modeling tax systems.

3.2 Aggregate Uncertainty and Undetectability

Recall how aggregate uncertainty has been formalized by Assumptions 3 and 4. For any $w \in W$, the "share" of individuals with productivity parameter w is commonly known. By contrast, the "share" of individuals with a taste parameter θ among those with productivity w, is not commonly known for all $w \in W$. Finally, those properties of the joint distribution of taste and skill parameters that are commonly known, determine the structure of the

---

[15] A similar reasoning can be found in Boylan (1998).

set $\mathcal{D}$.

The *undetectability* requirement in the above definition of a *C-IC* tax system precludes the formation of coalitions which induce an announced distribution of characteristics that does not belong to $\mathcal{D}$. Implicitly it is thus assumed, that the mechanism designer can effectively deter those collective manipulations for which it becomes obvious that some set of agents must have been deviating from the truth.[16]

The difficulty of achieving *C-IC* depends to a large extent on the assumptions on the feasible set $\mathcal{D}$ and the mechanism designer's ability to detect collective manipulations. To illustrate this, the example in subsections 2.1 and 2.4 is discussed once more.

3.3   The Example continued

Again consider the example discussed in subsections 2.1 and 2.4. Recall that it is assumed to be commonly known, that the average taste parameter $p\theta_H + (1-p)\theta_L$, is the same on every subset of W. Aggregate uncertainty stems only from the fact that p itself is an unknown parameter.

One can take the view that if the empirical taste and the empirical skill distribution satisfy this property of independence almost surely, then basically any collective manipulation is detectable. Whenever agents from a particular part of the skill distribution form a manipulating coalition – while all other agents stick to the truth – this induces an announced distribution of characteristics which is inconsistent with the commonly known fact that the average taste level is the same on every subset of W. Consequently, an undetectable manipulation has to be such that the average taste level is affected on every subinterval of W in the same way. This basically requires that the whole set of agents I is willing to undertake a collective manipulation. The only coalition which might potentially undermine an allocation is thus the so called *grand coalition* consisting of all agents. This is a perfectly consistent view on *undetectability*. It is formalized below as a undetectability in the *strict* sense.

With the following alternative notion of *undetectability* the mechanism cannot impose as much discipline on potential coalitions. As the realization of taste parameters is governed by an *i.i.d.* process of random variables, it may happen that the average taste level is different for different subintervals of W. Even though such an event has probability zero it is not excluded from the support of the stochastic process $\{\tilde{\theta}_w\}_{w \in W}$. This distinction between supported outcomes and those which arise with strictly positive probability allows for two different versions of the *undetectability* requirement:

---

[16]Note that even if a manipulation becomes apparent, the manipulating individuals are not yet identified. The above definition hence implicitly relies on the assumption, that the mechanism may punish *all* individuals harshly in response to an obvious collective lie. This in turn implies, that no coalition will consider such a collective plan.

Definition 5 Consider the application specified in subsections 2.1 and 2.4.

i) A collective manipulation is *weakly undetectable* if the induced skill distribution is given by F.

ii) A collective manipulation is *strictly undetectable* if the induced skill distribution is given by F and the induced taste distribution is such that average taste level is the same on every subinterval of the skill distribution.

Below, in section 5, the set of allocations which are not only *I-IC* but also *C-IC* is characterized for this environment. As will become clear, which version of *undetectability* is used, has a huge impact on the set of admissible allocations.

3.4 Related Concepts of Coalition-Proofness

The requirement of *C-IC* uses the notion of a *coalition-proof Nash equilibrium* that has been developed by Bernheim et al. (1986). These authors propose a refinement of the Nash equilibrium concept for games of complete information. As in this paper, the incentives coalition members face individually must not conflict with the action profile used by the coalition and, moreover, a conceivable collective manipulation has to be such that it does not provoke the formation of a further subcoalitions that depart from the initial coalitional plan (where a potentially deviating subcoalition would again have to meet these stability requirements.)

To relate their solution concept for games of complete information to the setting of this paper, the requirement of *C-IC* can be interpreted as follows. Suppose that, for some reason, the actual state of the economy D is commonly known among all individuals and that the mechanism designer is the only uninformed party. Still, the mechanism designer uses the revelation game to learn the actual state of the economy and to choose the level of public good provision $Q(D)$ and the menu of private goods bundles $B(D)$. The revelation game has thus become a game of complete information.[17] Moreover, each $D \in \mathcal{D}$ gives rise to a different complete information game. With this interpretation the requirement of *C-IC* can be stated as follows. *C-IC* holds if and only if in each such complete information game truth-telling is a stable best response for each coalition, given that all individuals outside the coalition tell the truth.

The insistence on stability of coalitions with respect to the formation of sub-coalitions distinguishes the present paper from some recent contributions to

---

[17]Moore (1992) provides a survey of implementation problems in environments with complete information.

the literature on mechanism design problems under the possibility of coalition formation. In a series of papers Laffont and Martimort (1997, 1999, 2000) incorporate a sequential Bayesian game of coalition formation into a mechanism design problem. These authors however only consider collective manipulations by the grand coalition. By contrast, Demange and Guesnerie (2001) allow for the formation of coalitions smaller than the grand coalition. As Laffont and Martimort they do not require stability with respect to the formation of subcoalitions. Instead they are concerned with concepts of the core in games of incomplete information without aggregate uncertainty.

4 Separability

In this section the set of allocations that are feasible, *I-IC* and *C-IC* is analyzed under the assumption that the utility function U is additively separable. This allows to establish a property which proves very useful in applications: The different incentive concerns can be separated. The requirement of *I-IC* deals with the resolution of pure assignment uncertainty in the profile of skill parameters; i.e. it allows to solve the screening problem of identifying individual skill levels within a given cross-section distribution F. The postulate of *C-IC* is concerned with problem of information aggregation which arises due to the aggregate uncertainty in the joint distribution of skill and taste parameters. It turns out that, in order to ensure *C-IC*, it suffices to eliminate incentives for a collective manipulation of taste parameters. There is no need to worry about collective manipulations of reported skill parameters.

Definition 6 A *utility allocation* is a mapping $\tilde{U} : (D, \gamma) \mapsto \tilde{U}(D, \gamma)$. A utility allocation $\tilde{U}$ is said to be *implementable* if there exists an anonymous allocation $[Q, A]$, which is feasible, *I-IC* and *C-IC* and such that:

$$\forall D, \forall \gamma : \quad \tilde{U}(D, \gamma) \ = \ U(Q(D), A(D, \gamma), \gamma) \ .$$

It will prove helpful to have an own terminology for coalitional manipulations which are based on a false report of taste parameters but which are truthful with respect to the reported skill parameters. A typical message profile of a manipulating coalition J which is such that, $\forall j \in J$, the reported skill parameter $\hat{w}_j$ is equal to the true skill parameter w_j is henceforth called a *partial taste manipulation* and denoted by $\hat{\gamma}_J^p$. To emphasize that some *manipulation* $\hat{\gamma}_J$ is not *partial*, I write $\hat{\gamma}_J = [\hat{w}_J, \hat{\theta}_J]$ with $\hat{w}_J \neq w_J$.

Definition 7

    i) A coalition J is said to possess a *partial taste manipulation* if there exists $D \in \mathcal{D}$ and an *undetectable* partial manipulation $\hat{\gamma}_J^p$ that induces a state perception $\hat{D}(\hat{\gamma}_J^p, D)$, which makes all members of J strictly better off relative to D (*unanimity*), prescribes for all its members taste announcements which are individually a best response in conjunction with a truthful skill announcement under state perception $\hat{D}(\hat{\gamma}_J^p, D)$ (*individual stability*), and is not threatened by further partial taste manipulations of subcoalitions, which satisfy all these requirements (*collective stability*).

    An allocation is said to have the *collective revelation of taste (C-RT)* property if there does not a exist a coalition with a partial taste manipulation.

    ii) A utility allocation is *partially implementable* if there exists an anonymous allocation $[Q, A]$, which is feasible, *I-IC*, has the *C-RT* property and is such that:

$$\forall D, \forall \gamma : \quad \tilde{U}(D, \gamma) \;=\; U(Q(D), A(D, \gamma), \gamma) \,.$$

Obviously, if an allocation is *C-IC*, then it has also the *C-RT* property. As a consequence, the set of implementable utility allocations is a subset of the set of partially implementable allocations. The following Lemma shows that the converse inclusion holds true as well. Hence, it justifies an analysis of allocations which possess only the *C-RT* property.

Lemma 1 Under assumptions 1 - 4, the set of implementable utility allocations is equal to the set of partially implementable utility allocations.

The proof is based on the observation that under aggregate stability with respect to the distribution of skill parameters, any conceivable undetectable collective manipulation which involves both taste and skill parameters can be mimicked by a partial manipulation which involves only reported taste parameters. Intuitively, any manipulation has to be undetectable. Hence, whenever a subset of agents J manipulates via some $\hat{\gamma}_J = [\hat{w}_J, \hat{\theta}_J]$ with $\hat{w}_J \neq w_J$, this manipulation has to be such that the resulting distribution of announcements $\hat{D}$ has a marginal skill distribution which is equal to F. But this implies that coalition J can induce $\hat{D}$ as well by a suitably chosen partial taste manipulation. As a consequence, it suffices to exclude the possibility of partial taste manipulations in order to establish the *C-IC* property.

The results in Lemmas 2 and 1 imply that, under assumptions 1-4, attention can be restricted to the set of feasible allocations which satisfy *I-RP*, *NDT* and *C-RT*. These observations are summarized in the following Theorem.

Proposition 3 Under assumptions 1-4, any implementable utility allocation is also implementable via an allocation $[Q, A]$ that satisfies the following properties: *I-RP*, *NDT*, *C-RT* and *feasibility*.

In Proposition 2 it was established that an allocation that is decentralizable by a tax system needs to be feasible, must not discriminate between individuals who differ in their views on public goods but are otherwise identical (*NDT*-property), and, finally, has to ensure that individuals are willing to reveal their earning ability (*I-RP*-property). These are the constraints that are familiar from the theory of optimal income taxation.

According to Proposition 3 one set of constraints has to be added if there is uncertainty about the distribution of preferences in the economy and individuals can form coalitions in order to manipulate the tax system and the provision rule for public goods. These *C-RT* constraints ensure that no group of agents gains from a collective lie on their taste parameter.

Form a more general perspective this result can be interpreted as follows. The constraints analyzed in the theory of optimal taxation are concerned with individual behavior under a given tax system. These constraints remain relevant in the present setting with aggregate uncertainty. In addition, there is a set of *"political economy constraints"* which addresses the ability of individuals to form interest groups if, prior to a decision on public good provision, information on preferences has to be acquired.

The *C-RT* constraints may seem rather opaque at the present level of abstraction. It is not obvious how to represent them by a well-defined set of constraints that could, for instance, be included in an exercise of solving for an optimal constrained efficient allocation. To illustrate the impact of the *C-RT* property the next section returns to the example discussed in subsections 2.1, 2.4 and 3.3.

5   The Example continued

This section returns to the application already discussed in subsections 2.1, 2.4 and 3.3. For this environment the set of allocations that are *feasible*, *I-IC* and *C-IC* is explicitly characterized in the following. Finally, I briefly discuss the properties of *optimal* allocations that meet all these criteria.

By Proposition 3 attention is restricted to allocations $[Q, t]$ that satisfy *I-RP*, *NDT*, *feasibility* and the *C-RT* property. As has already been observed in subsection 2.4, the first three requirements are equivalent to the payment scheme t being such that the cost of public good provision is shared equally among all individuals, for every state p of the economy. Consequently, the remaining task is to characterize the provision rules $Q : p \mapsto Q(p)$ that yield the *C-RT* property under such a payment scheme. To achieve this, the two

versions of the *undetectability* requirement, that have been introduced in Definition 5 have to be distinguished.

5.1 Undetectability in the strict sense

Recall that under *undetectability* in the strict sense the mechanism designer infers the actual value of the parameter p from a profile of reports $\{\hat{\gamma}_i\}_{i \in I} = \{(\hat{\theta}_i, \hat{w}_i)\}_{i \in I}$ and is able to deter any manipulation such that the reported average taste level is different on different subintervals of W.

Proposition 4 Suppose a manipulation is called undetectable if it is *strictly undetectable* in the sense of Definition 5. Any pair $[Q, t]$ that satisfies *equal cost sharing* also satisfies the *C-RT* property if there do not exist p and p' such that,

$$\forall w, \forall \theta : \quad V(p, \theta, w) \quad > \quad V(p', \theta, w) . \tag{2}$$

Under *undetectability* in the strict sense basically any provision rule $Q : p \mapsto Q(p)$ is implementable if accompanied by equal cost sharing. The only additional restriction imposed by *C-RT* is that there must not exist a state of the world p such that all individuals unanimously agree that there exists a preferred outcome of the revelation game.[18] This implies in particular, that provision rule $Q^*(p)$ which maximizes EW pointwise, is implementable.

5.2 Undetectability in the weak sense

Under *undetectability* in the weak sense, an inconsistency of average taste levels on different subintervals of W is a possible event. The following assumption specifies how the mechanism designer responds to such an event.

Assumption 6 The mechanism designer's perception of p is given by $\mu(\{i \mid \hat{\theta}_i = \theta_H\})$, where $\hat{\theta}_i$ is the taste announcement of individual i in the revelation game.[19]

According to Assumption 6, the mechanism designer takes the population share of individuals with a high taste parameter as the state of the economy.

---

[18]This is a sufficient condition. (2) implies that there does not exist a partial taste manipulation for the *grand coalition* of all agents which satisfies *Undetectability, Unanimity* and *Individual Stability*.

[19]Measurability of the set $\{i \mid \hat{\theta}_i = \theta_H\}$ is again ensured with reference to Al-Najjar (2004). In his model of a large economy I is a countable set of infinitely many individuals and the set $\{i \mid \hat{\theta}_i = \theta_H\}$ is measurable with respect to an appropriate generalization of the counting measure.

Put differently, the mechanism designer chooses provision level $Q(p)$ whenever he observes that $\mu(\{i \mid \hat{\theta}_i = \theta_H\})$ is equal to p.[20] Consequently, under assumption 6 there is an obvious channel along which a coalition might manipulate an allocation. Any partial taste manipulation that affects the population share of individuals who announce a high taste parameter has an effect on the level of public good provision. In particular, this implies that any partial taste manipulation that affects the average taste level on some subinterval of W becomes effective. The following proposition derives the implications of this property for the set of implementable allocations.

Proposition 5 Suppose that assumption 6 applies. Consider an allocation $[Q,t]$ with equal cost sharing. $[Q,t]$ has the *C-RT* property for any minimal coalition size ϵ if and only if the following properties are satisfied.

i) Q is a non-decreasing function of p.

ii) $V(p, \theta_L, \bar{w})$ is non-increasing and $V(p, \theta_H, \bar{w})$ is non-decreasing in p.

The *"if-part"* in the proposition follows from the observations that, under a non-decreasing provision rule, $V(p, \theta_L, \bar{w})$ is non-increasing in p only if an individual with effective valuation $\theta_L \bar{w}$ always desires a small provision level over a large provision level. This implies that the same is true for any individual with an effective valuation $\theta_L w \leq \theta_L \bar{w}$. As a consequence, no individual with a low taste realization is willing to join a manipulating coalition that attempts to achieve a larger perception of the average taste parameter p, or, equivalently, a larger quantity of the public good. Analogously one shows that no individual with a high taste realization wants to achieve a smaller perception of p. Consequently, even with the opportunity to undertake manipulative collective actions, individuals cannot to better than to reveal their taste parameter.

The proof of the *"only if-part"* is based on the observation that whenever property i) or property ii) is violated, then there exists some small ϵ such that the *C-RT* property fails. For instance, as shown in the appendix, if there exist p' and p with $p' - p = \epsilon$ and $Q(p') < Q(p)$, then there exists a small coalition of individuals with a low taste realization that tries to induce the outcome $Q(p')$ if the true state is p, or a coalition of high taste individuals that aims at $Q(p)$ if the true state is p'. Hence, the *C-RT* property holds for any small ϵ only if properties i) and ii) are fulfilled.

Using Proposition 5 it is easily verified that Q^*, the welfare maximizing

---

[20]The provision rule Q can thus be viewed as resulting from a voting procedure. To see this, interpret a high (low) taste announcement as a vote in favor of a large (small) level of public good provision. In this sense, the provision rule $p \mapsto Q(p)$ specifies a provision level for each conceivable vote distribution; see Bierbrauer and Sahm (2006) for more details.

provision rule under equal cost sharing, is not part of an implementable allocation. For instance, as has already been discussed in subsection 2.4, there exists a range of small values of p such that the indirect utility function under Q^*, $V^*(p, \theta_L, \bar{w})$, is strictly increasing in p.

To see that the requirement of C-RT can affect the properties of the optimal provision rule very drastically, suppose that the parameters of the model are such that $\theta_L \bar{w} > \theta_H \underline{w}$; i.e. an individual with a very low skill level and a high taste parameter has an effective valuation of the public good which is smaller than the one of an individual with a very high skill level and a low taste parameter.[21] It is easily verified that only a constant provision rule satisfies the C-RT property.[22] Put differently, the C-RT requirement is met only if information aggregation plays no role. In this case, the optimal provision rule is characterized by the condition $Q(p) = Q^*(\frac{1}{2})$ for all p.

6   Concluding Remarks

This paper has introduced the notion of a *collective incentive compatible tax system*. Such a tax system acquires private information of individuals on their preferences in a way that does not leave scope for collective manipulations.

The theory of optimal taxation has traditionally been concerned with an optimal use of a given set of tax instruments that takes the behavioral responses of individuals to changes in the tax system into account. The additional requirement of *collective incentive compatibility* captures that individuals may also respond collectively to the tax policy they are confronted with. One might thus hope that the study of *collective incentive compatible tax system* provides a link between the normative approach in the theory of optimal taxation and the field of political economics.

However, the definition of a collectively incentive compatible allocation as such is rather involved and does not immediately provide a well defined set of constraints. The approach adopted in this paper (and as well in Bierbrauer (2005)) has thus been to look at a more structured environment in which this requirement can be represented by a tractable set of inequality constraints.

The question whether there is an alternative to this problem-specific approach which yields a more general characterization of *individually* and *collectively incentive compatible* allocations is left to future research.

---

[21]For the optimal provision rule under the assumption $\theta_L \underline{\lessgtr} \theta_H \underline{w}$, the reader is referred to Bierbrauer and Sahm (2006). Even though that paper is concerned voting mechanisms it arrives at the same characterization of implementable provision rules as Proposition 5.

[22]To see this, let $p' > p$. Property i) in Proposition 5 implies $Q(p') \geq Q(p)$. Property ii) implies that $\theta_H \underline{w}(Q(p') - Q(p)) \geq K(Q(p')) - K(Q(p)) \geq \theta_L \bar{w}$. With $\theta_L \bar{w} > \theta_H \underline{w}$ the only way to satisfy all these conditions is to have $Q(p) = Q(p')$.

A Appendix

Statement and Proof of Revelation Principle

An *anonymous mechanism* M is a game form consisting of a message space R, a provision rule for the public good and a menu of consumption-income combinations. To describe these functions denote by Δ_R the set of cumulative distribution functions (cdfs) on R and denote a typical element of Δ_R by ρ. An anonymous mechanism is defined by the mappings:

$$Q^M : \Delta_R \to \mathbb{R}_+, \ \rho \mapsto Q^M(\rho) \ ,$$

$$A^M : \Delta_R \times R \to \mathbb{R}_+^2, \ (\rho, r) \mapsto A^M(\rho, r) \ .$$

A *direct anonymous mechanism* $\bar{M}$ is an anonymous mechanism which satisfies $R = \Gamma$ and is summarized by the functions

$$Q^{\bar{M}} : \Delta_\Gamma \to \mathbb{R}_+, \ D \mapsto Q^{\bar{M}}(D) \ ,$$

$$A^{\bar{M}} : \Delta_\Gamma \times \Gamma \to \mathbb{R}_+^2, \ (D, \gamma) \mapsto A^{\bar{M}}(D, \gamma) \ .$$

Note that the domain of a *direct anonymous mechanism* does not coincide with the one of an anonymous allocation defined in the body of the text. The reason is that an anonymous allocation specifies the level of public good provision Q and the menu of consumption-income pairs A only for cdfs which belong to the feasible set $\mathcal{D}$. By contrast, a *direct anonymous mechanism* specifies an outcome of the game for each distinguishable action profile, that is, for each distribution of announcements in Δ_Γ.

Consider the game induced by an anonymous mechanism M. A *strategy* s for an agent assigns a report to each possible value of individual characteristics. Formally:

$$s : \Gamma \to R : r = s(\gamma) \ .$$

Denote the set of possible strategies by S.

The game induced by anonymous mechanism M has an *equilibrium in dominant strategies* if there exists a mapping s^* such that $\forall \gamma$, $\forall \rho \in \Delta_R$ and $\forall s \in S$:

$$U(Q^M(\rho), A^M(\rho, s^*(\gamma)), \gamma) \geq U(Q^M(\rho), A^M(\rho, s(\gamma)), \gamma) \ .$$

In words: Each type γ has a best response $s^*(\gamma)$, which applies independently of the behavior of others, i.e. which is optimal for all $\rho \in \Delta_R$.

An anonymous mechanism M *implements* an anonymous allocation *in dominant strategies* if the game induced by M has an equilibrium in dominant strategy s^* which satisfies $\forall \gamma$ and $\forall D \in \mathcal{D}$:

$$Q(D) = Q^M(\rho^*(D)) \quad \text{and} \quad A(D, \gamma) = A^M(\rho^*(D), s^*(\gamma)) \ ,$$

where $\rho^*(D)$ is the distribution of reports generated by s^* if the state of the economy is D. Put differently $\rho^*(D)$ is the distribution on R induced by the message profile $\{s^*(\gamma_i)\}_{i \in I}$ if the cdf that corresponds to the profile of characteristics in $\{\gamma_i\}_{i \in I}$ is D.

Consider the game induced by a direct anonymous mechanism $\bar{M}$. *Truth-telling* is a strategy defined by $s(\gamma) = \gamma$ for all $\gamma \in \Gamma$. Truth-telling by all agents is an equilibrium in dominant strategies provided that $\forall \gamma, \forall \hat{\gamma}$ and $\forall D \in \Delta_\Gamma$:

$$U(Q^{\bar{M}}(D), A^{\bar{M}}(D, \gamma), \gamma) \geq U(Q^{\bar{M}}(D), A^{\bar{M}}(D, \hat{\gamma}), \gamma) . \tag{3}$$

An anonymous allocation is *truthfully implementable* in dominant strategies if there exists a direct anonymous mechanism $\bar{M}$ that implements it such that truth-telling is a dominant strategy; i.e. truthful implementation requires that truth-telling by all agents is an equilibrium in dominant strategies in the game induced by $\bar{M}$, and in addition $\forall \gamma$ and $\forall D \in \mathcal{D}$:

$$Q(D) = Q^{\bar{M}}(D) \quad \text{and} \quad A(D, \gamma) = A^{\bar{M}}(D, \gamma) . \tag{4}$$

Lemma 2 An anonymous allocation rule is *I-IC* if and only if it is truthfully implementable.

Proof The if-part follows from substituting the equations in (4) into the inequalities in (3), for $D \in \mathcal{D}$. This yields the definition of an *I-IC* allocation. To prove the only if-part, suppose that the pair $[Q, A]$ is an *I-IC* anonymous allocation rule. It has to be shown that there exists a direct anonymous mechanism $[Q^{\bar{M}}, A^{\bar{M}}]$ which implements $[Q, A]$. This direct anonymous mechanism has to be such that the incentive structure is preserved, i.e. such that truth-telling is a dominant strategy. It can for instance be constructed as follows. For all $\gamma \in \Gamma$ and all $D \in \mathcal{D}$ choose $[Q^{\bar{M}}, A^{\bar{M}}]$ such that (4) holds. For all $D \in \Delta_\Gamma \backslash \mathcal{D}$ and $\gamma \in \Gamma$, let $Q(D) = $ constant and $A(D, \gamma) = $ constant. ∎

Proposition 6 (Revelation Principle) An anonymous allocation is implementable if and only if it is truthfully implementable.

Proof The if-part is trivial. Suppose $[Q, A]$ is implementable by some mechanism M. Then there exists a function s^* such that $\forall \gamma, \forall \hat{\gamma}, \forall D \in \mathcal{D}$:

$$U(Q^M(\rho^*(D)), A^M(\rho^*(D), s^*(\gamma)), \gamma) \geq U(Q^M(\rho^*(D)), A^M(\rho^*(D), s^*(\hat{\gamma})), \gamma) .$$

(In words: In a dominant strategy equilibrium, the following has to be true. The actions prescribed by the equilibrium strategy s^* are such that no

type wants to deviate to an action prescribed for another type, taking the distribution over equilibrium actions as given.) and such that $\forall \gamma, \forall D \in \mathcal{D}$:

$$Q(D) = Q^M(\rho^*(D)) \quad \text{and} \quad A(D,\gamma) = A^M(\rho^*(D), s^*(\gamma)) \,.$$

Combining those statements yields the definition of truthful implementability or equivalently of an *I-IC* anonymous allocation.

■

Proof of Taxation Principle in Example 2.
"⟸": Consider a feasible anonymous allocation. Suppose it is an income tax but not *I-IC*. Then there exist γ, $\hat{\gamma}$ and D such that

$$U(Q(D), A(D,\gamma), \gamma) < U(Q(D), A(D,\hat{\gamma}), \gamma) \,.$$

Using that for all γ, $A(D,\gamma) = [Y(D,\gamma) - T(D,Y(D,\gamma)), Y(D,\gamma)]$, this is equivalent to

$$U(Q(D), Y(D,\gamma) - T(D,Y(D,\gamma)), Y(D,\gamma), \gamma)$$

$$< U(Q(D), Y(D,\hat{\gamma}) - T(D,Y(D,\hat{\gamma})), Y(D,\hat{\gamma}), \gamma) \,.$$

But this contradicts that $\forall D, \forall \gamma$:

$$Y(D,\gamma) \quad \in \quad \arg\max_Y U(Q(D), Y - T(D,Y), Y, \gamma)$$

"⟹": Consider a feasible and *I-IC* allocation and construct T as follows:

  i) For any x such that there is D and γ with $Y(D,\gamma) = x$ define $T(D,x)$ by the equation[23]

$$T(D,x) \quad = \quad Y(D,\gamma) - C(D,\gamma) \,.$$

  Obviously, this choice ensures that under T, consumption equals after tax income and that budget balance holds.

  ii) For all other levels of Y set $T(D,x) = x$.[24]

---

[23]Note that his equation uniquely determines $T(D,x)$. If not, one had, for given D, different consumption levels corresponding to the same income requirement; hence a contradiction to individual incentive compatibility, assuming monotonicity of preferences.

[24]It is implicitly assumed that, for any agent, zero consumption implies a utility level of $-\infty$ and that hence the corresponding Y is never chosen, whenever there is an alternative with positive consumption available.

Now suppose this function T does not satisfy the property that $\forall D, \forall \gamma$:

$$Y(D, \gamma) \in \arg\max_Y U(Q(D), Y - T(D, Y), Y, \gamma)$$

Then there exist $\gamma, \hat{\gamma}$ and D such that

$$U(Q(D), Y(D, \gamma) - T(D, Y(D, \gamma)), Y(D, \gamma), \gamma)$$

$$< U(Q(D), Y(D, \hat{\gamma}) - T(D, Y(D, \hat{\gamma})), Y(D, \hat{\gamma}), \gamma) \, .$$

or using that for all γ, $A(D, \gamma) = [Y(D, \gamma) - T(D, Y(D, \gamma)), Y(D, \gamma)]$,

$$U(Q(D), A(D, \gamma), \gamma) < U[Q(D), A(D, \hat{\gamma}), \gamma] \, .$$

This contradicts *I-IC*.

$\blacksquare$

Proof of Proposition 2. To proof the only if-part note that, because preferences satisfy Assumption 1, the *NDT-U* property is an implication of *I-IC*. Obviously *I-RP* is also an implication of *I-IC*. To prove the if-part, suppose an allocation rule, such that the *NDT-U* and the *I-RP* property hold, is not *I-IC*. Then there exist (θ, w) and $(\hat{\theta}, \hat{w})$ and D such that $u(A(D, \theta, w), w) < u(A(D, \hat{\theta}, \hat{w}), w)$. Using *NDT-U* and *I-RP* one has:

$$u(A(D, \hat{\theta}, \hat{w}), w) = u(A(D, \theta, \hat{w}), w) \leq u(A(D, \theta, w), w) \, .$$

Hence, a contradiction.

$\blacksquare$

Proof of Lemma 1. It has to be shown that any partially implementable utility allocation is implementable. Suppose to the contrary that there exists a partially implementable utility allocation $\tilde{U}$, which is not implementable.

  i) Denote by $[Q, A]$ the feasible, *I-IC* and *C-RT* allocation which partially implements $\tilde{U}$. By hypothesis $\tilde{U}$ is not implementable. Hence, there must exist D and a coalition J and a manipulation $\hat{\gamma}_J = [\hat{w}_J, \hat{\theta}_J]$ with $\hat{w}_J \neq w_J$ such that, by *Undetectability*, $\hat{D}(\hat{\gamma}_J, D) \in \mathcal{D}$ and,

$\forall i \in J$, by *Individual Stability* and *Unanimity*

$$v(Q(\hat{D}(\hat{\gamma}_J, D)), \theta_i) + u(A(\hat{D}(\hat{\gamma}_J, D), \hat{\theta}_i, \hat{w}_i), w_i)$$

$$= v(Q(\hat{D}(\hat{\gamma}_J, D)), \theta_i) + u(A(\hat{D}(\hat{\gamma}_J, D), \theta_i, w_i), w_i) \tag{5}$$

$$> v(Q(D), \theta_i) + u(A(D, \theta_i, w_i), w_i)$$

and such that *collective stability holds.*

ii) *Claim.* The coalition J can induce the announced distribution $\hat{D}(\hat{\gamma}_J, D)$ also via some partial taste manipulation $\hat{\gamma}_J^p$.
Proof. There is aggregate stability with respect to the marginal distribution of skill parameters. Hence, any undetectable manipulation $\hat{\gamma}_J$ with $\hat{w}_J \neq w_J$ has to be consistent with the commonly known skill distribution F. The manipulation of J presumes that all individuals not in J reveal their characteristics truthfully. Hence, to be undetectable, $\hat{\gamma}_J$ has to be such that the distribution of skill announcements within coalition J is equal to the true skill distribution within coalition J. But this implies that the outcome achieved via $\hat{\gamma}_J$ is also induced if all members of J reveal their skill parameter truthfully and choose a suitable profile of announced taste parameters. I.e. for given $\hat{\gamma}_J$ with $\hat{w}_J \neq w_J$, there exists $\hat{\gamma}_J^p$ with $\hat{w}_J = w_J$ such that $\hat{D}(\hat{\gamma}_J, D) = \hat{D}(\hat{\gamma}_J^p, D)$.

iii) *Claim.* The partial taste manipulation $\hat{\gamma}_J^p$ defined with reference to $\hat{\gamma}_J$ in ii) satisfies *Individual Stability* and *Unanimity*.
Proof. $\hat{\gamma}_J^p$ is a partial taste manipulation. Under the separability assumption 1, *I-IC* implies the *NDT* property. Hence, any partial manipulation satisfies *Individual Stability*. *Unanimity* follows from $\hat{D}(\hat{\gamma}_J, D) = \hat{D}(\hat{\gamma}_J^p, D)$ and the inequality in (5).

iv) If $\hat{\gamma}_J^p$ was collectively stable with respect to partial taste manipulations by subcoalitions of J, then this would contradict, the *C-RT* property of utility allocation $\tilde{U}$. Hence, I assume in the following that $\hat{\gamma}_J^p$ is not collectively stable with respect to partial taste manipulations by subcoalitions of J. I.e. if the true distribution of characteristics in the economy is D and coalition J has induced the announced distribution $\hat{D}(\hat{\gamma}_J, D)$, then there exists a subcoalition J' of J with a partial taste manipulation $\tilde{\gamma}_{J'}^p \neq \hat{\gamma}_{J'}^p$ which induces a state perception $\hat{D}(\tilde{\gamma}_{J'}^p, \hat{\gamma}_{J \setminus J'}^p, D) \in \mathcal{D}$ (*Undetectability*), which is strictly preferred by all members of J' relative to $\hat{D}(\hat{\gamma}_J, D)$ (*Unanimity*), is *individually stable* and does not provoke partial taste manipulations by further subcoalitions (*collective stability*).

v) *Claim.* It has to be true that the partial taste manipulation $\tilde{\gamma}^p_{J'}$ by subcoalition J' characterized in iv) is not collectively stable with respect to all manipulations $\bar{\gamma}_{J''} \neq \tilde{\gamma}_{J''}$ with $\bar{w}_J \neq w_J$ by subcoalitions J'' of J'.

Proof. Suppose otherwise. Then this partial taste manipulation $\tilde{\gamma}^p_{J'}$ could be used by the set J' to manipulate the initial manipulation of allocation $[Q, A]$ by coalition J via $\hat{\gamma}_J$ in step i), thereby contradicting the *collective stability* of this manipulation.

vi) The reasoning established so far has a recursive structure: The starting point was in step i) an allocation $[Q, A]$, which is not vulnerable by partial taste manipulations but by a joint manipulations $\hat{\gamma}_J$ of both taste and skill parameters. In steps ii)-v) it has been shown that this implies the existence of a subcoalition J' of J which possess a partial taste manipulation $\tilde{\gamma}^p_{J'}$ which does not provoke further partial taste manipulations by subcoalitions of J' but further joint manipulations of both taste and skill parameters by subcoalitions of J'.

Now the reasoning in steps i) - v) can be applied again to show that this implies the existence of a subcoalition J'' of J' which possess a partial taste manipulation but provokes further joint manipulations of both taste and skill parameters by subcoalitions of J'' etc.

However, as a consequence of definition 3, any chain of successive formation of subcoalitions has a finite length. Hence, after a finite number of repeated applications of the reasoning in steps i) - v) one ends up with a situation in which a subcoalition J^x of minimal size – that is, J^x possesses no further subcoalitions – possess a joint manipulations of both taste and skill parameters but not a partial taste manipulation. A last application of steps ii) and iii) then yields a contradiction.

∎

Proof of Proposition 4. If (2) holds, then there is no coalition which is willing to affect the average taste level on all subintervals of W. Any coalition which affects the average taste level only on some subintervals of W is detected.

∎

Proof of Proposition 5. The proof follows from Lemmas 3 – 6 below.

Lemma 3 Suppose that assumption 6 applies. Let the minimal coalition size ϵ be close to zero. If an allocation $[Q, t]$ with equal cost sharing satisfies the *C-RT* property, then $Q(p') \geq Q(p)$ for any pair $p', p \in (0, 1)$ with $\epsilon < p' - p \leq 2\epsilon$.

Proof Consider a pair $p', p \in (0,1)$ that satisfies $\epsilon < p' - p \leq 2\epsilon$. If ϵ is sufficiently small, then there exists some skill interval $[w_1, w_2] \subset W$, with $\theta_H w_1 > \theta_L w_2$ and the following property: under assumption 5, for all p, almost surly, there exist coalitions $J_L \subset I$ and $J_H \subset I$ such that:

i) All members of J_L and J_H have a skill parameter within $[w_1, w_2]$. Moreover, for all $i \in J_L$, $\theta_i = \theta_L$ and for all $i \in J_H$, $\theta_i = \theta_H$.

ii) Both coalitions are of equal size, possess no subcoalitions and satisfy

$$p' - p \;=\; \mu(J_L) = \mu(J_H) \,.$$

Suppose first that the true average taste parameter is given by p. As $[Q, t]$ is *C-RT* there exists $i \in J_L$ such that

$$\theta_L Q(p) - \frac{K(Q(p))}{w_i} \;\geq\; \theta_L Q(p') - \frac{K(Q(p'))}{w_i} \,. \tag{6}$$

Suppose to the contrary that there does not exist such an $i \in J_L$. Then if all individuals in J_L announce a high taste parameter this yields a partial taste manipulation which satisfies *weak undetectability, unanimity, individual stability*, because due to the *NDT* property individuals are willing to announce any taste parameter, and *collective stability*, as J_L has no subcoalition. Now suppose that the true aggregate taste level is given by p'. Analogously, there exists $j \in J_H$ such that

$$\theta_H Q(p') - \frac{K(Q(p'))}{w_j} \;\geq\; \theta_H Q(p) - \frac{K(Q(p))}{w_j} \,. \tag{7}$$

Combining the inequalities (6) and (7) yields:

$$(\theta_H w_j - \theta_L w_i)(Q(p') - Q(p)) \;\geq\; 0 \,.$$

By construction, for all $w_i, w_j \in [w_1, w_2]$, $\theta_H w_j - \theta_L w_i > 0$. Hence, it has to be true that $Q(p') \geq Q(p)$. ∎

Lemma 4 Suppose that assumption 6 applies. Let the minimal coalition size ϵ be close to zero. If an allocation $[Q, t]$ with equal cost sharing satisfies *C-RT*, then: for all $w \in W$, and for any pair $p', p \in (0,1)$ with $\epsilon < p' - p \leq 2\epsilon$, $V(p', \theta_L, w) \leq V(p, \theta_L, w)$ and $V(p', \theta_H, w) \geq V(p, \theta_H, w)$.

Proof Without loss of generality, suppose that there exist p' and p with $\epsilon < p' - p \leq 2\epsilon$ and $w \in W$ such that $V(p, \theta_H, w) > V(p', \theta_H, w)$. As ϵ is small, there exists an interval $[w_1, w_2] \subset W$ with $w_1 \leq w \leq w_2$ and

$w_1 < w_2$ and a coalition J_H with the following properties: For all $i \in J_H$, $w_i \in [w_1, w_2]$ and $\theta^i = \theta_H$ and, moreover, $\mu(J_H) = p' - p$. Obviously, if the true average taste parameter equals p, this coalition possesses a partial taste manipulation. This contradicts the C-RT property of $[Q, t]$.

∎

As a consequence of Lemmas 3 and 4, whenever the provision rule Q is such that for some pair p', p with $p < p'$ one has $Q(p) > Q(p')$ or, $V(p', \theta_L, \bar{w}) > V(p, \theta_L, \bar{w})$ or $V(p', \theta_H, \underline{w}) < V(p, \theta_H, \underline{w})$, then there exists a value for the minimal coalition size ϵ such that the C-RT property is violated.

Lemma 5 If for any $w \in W$, $V(p, \theta_L, w)$ is non-increasing and $V(p, \theta_L, w)$ is non-decreasing in p then the C-RT property is implied, for any minimal coalition size $\epsilon > 0$.

Proof Suppose that $[Q, t]$ does not have the C-RT property. Then there exists a level of the true aggregate taste parameter p and a coalition J with a partial taste manipulation which induces an announced aggregate taste level of $p' \neq p$. Without loss of generality, assume that $p' > p$. Suppose that J contains an individual with a low taste parameter. Due to the *unanimity* property, this individual is made strictly better off by this partial taste manipulation. This contradicts the assumption that $V(p, \theta_L, \bar{w})$ is non-increasing in p. Now suppose that J contains only of individuals with a high taste parameter. If the true aggregate taste level is p and individuals in J misreport their taste parameter, this cannot induce an announced aggregate taste level exceeding p.

∎

Lemma 6 Suppose that assumption 6 applies. Consider an allocation $[Q, t]$ with equal cost sharing. Let Q be a non-decreasing function of p. Then, the C-RT property holds if $V(p, \theta_L, \bar{w})$ is non-increasing in p, and $V(p, \theta_H, \underline{w})$ is non-decreasing in p.

Proof If $V(p, \theta_L, \bar{w})$ is non-increasing in p, one has for all p and all p' with $p' \geq p$ that

$$\theta_L \bar{w}(Q(p') - Q(p)) \leq K(Q(p')) - K(Q(p)) .$$

As Q is non-decreasing in p, this implies that $\forall w \in W$,

$$\theta_L w(Q(p') - Q(p)) \leq K(Q(p')) - K(Q(p)) .$$

Hence, for all w, $V(p, \theta_L, w)$ is non-increasing in p. Analogously one shows that if $V(p, \theta_H, \underline{w})$ is non-decreasing in p, this implies that, for all w $V(p, \theta_H, w)$ is non-decreasing in p. Using Lemma 5 this establishes the *C-RT* property. ∎

References

Al-Najjar, N. (2004). Aggregation and the law of large numbers in large economies. *Games and Economic Behavior*, 47:1–35.

Atkinson, A. and Stern, N. (1974). Pigou, taxation and public goods. *Review of Economic Studies*, 41:119–128.

Bergemann, D. and Morris, S. (2005). Robust mechanism design. *Econometrica*, 73:1771–1813.

Bernheim, B., Peleg, B., and Whinston, M. (1986). Coalition-proof Nash equilibria I. concepts. *Journal of Economic Theory*, 42:1–12.

Bierbrauer, F. (2005). Optimal income taxation and public good provision in a two-class economy. Preprint 2005/25, Max Planck Institute for Research on Collective Goods, Bonn.

Bierbrauer, F. (2006). Distortionary taxation and the free-rider problem. Preprint 2006/6, Max Planck Institute for Research on Collective Goods, Bonn.

Bierbrauer, F. and Sahm, M. (2006). Informative voting and the Samuelson rule. Preprint 2006/18, Max Planck Institute for Research on Collective Goods, Bonn.

Boadway, R. and Keen, M. (1993). Public goods, self-selection and optimal income taxation. *International Economic Review*, 34:463–478.

Boylan, T. (1998). Coalition-proof implementation. *Journal of Economic Theory*, 82:132–143.

Clarke, E. (1971). Multipart pricing of public goods. *Public Choice*, 11:17–33.

Demange, G. and Guesnerie, R. (2001). On coalitional stability of anonymous interim mechanisms. *Economic Theory*, 18:367–389.

Gaube, T. (2000). When do distortionary taxes reduce the optimal supply of public goods. *Journal of Public Economics*, 76:151–180.

Gaube, T. (2005). Financing public goods with income taxation: Provision rules vs. provision levels. *International Tax and Public Finance*, 12:319–334.

Groves, T. (1973). Incentives in teams. *Econometrica*, 41:617–663.

Guesnerie, R. (1995). *A Contribution to the Pure Theory of Taxation*. Cambridge University Press.

Hammond, P. (1979). Straightforward individual incentive compatibility in large economies. *Review of Economic Studies*, 46:263–282.

Hellwig, M. (2003). Public-good provision with many participants. *Review of Economic Studies*, 70:589–614.

Hellwig, M. (2004). Optimal income taxation, public goods provision and public-sector pricing: A contribution to the foundations of public economics. Preprint 2004/14, Max Planck Institute for Research on Collective Goods, Bonn.

Hellwig, M. (2005). A utilitarian approach to the provision and pricing of excludable public goods. *Journal of Public Economics*, 89:1981–2003.

Kalai, E. (2004). Large robust games. *Econometrica*, 72:1631–1665.

Laffont, J. (1987). Incentives and the allocation of public goods. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*. North-Holland, Amsterdam.

Laffont, J. and Martimort, D. (1997). Collusion under asymmetric information. *Econometrica*, 65:875–911.

Laffont, J. and Martimort, D. (1999). Collusion-proof samuelson conditions for public goods. *Journal of Public Economic Theory*, 1:399–438.

Laffont, J. and Martimort, D. (2000). Mechanism design with collusion and correlation. *Econometrica*, 68:309–342.

Mas-Colell, A., Whinston, M., and Green, J. (1995). *Microeconomic Theory*. Oxford University Press, New York.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.

Moore, J. (1992). Implementation, contracts, and renegotiation in environments with complete information. In Laffont, J.-J., editor, *Advances in Economic Theory: Sixth World Congress, vol. I*. Cambridge, UK, Cambridge University Press.

Nava, M., Schroyen, F., and Marchand, M. (1996). Optimal fiscal and public expenditure policy in a two-class economy. *Journal of Public Economics*, 61:119–137.

Norman, P. (2004). Efficient mechanisms for public goods with use exclusion. *Review of Economic Studies*, 71:1163–1188.

Samuelson, P. (1954). The pure theory of public expenditure. *Review of Economics and Statistics*, 36:387–389.

Sandmo, A. (1998). Redistribution and the marginal cost of public funds. *Journal of Public Economics*, 70:365–382.

Wilson, J. (1991). Optimal public good provision with limited lump sum taxation. *American Economic Review*, 81:153–166.