

Bierbrauer, Felix

**Working Paper**

## Optimal income taxation and public goods provision in a large economy with aggregate uncertainty

CESifo Working Paper, No. 2701

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Bierbrauer, Felix (2009) : Optimal income taxation and public goods provision in a large economy with aggregate uncertainty, CESifo Working Paper, No. 2701, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/30683>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Optimal Income Taxation and Public Goods Provision in a Large Economy with Aggregate Uncertainty

FELIX BIERBRAUER

CESIFO WORKING PAPER NO. 2701

CATEGORY 1: PUBLIC FINANCE

JULY 2009

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Optimal Income Taxation and Public Goods Provision in a Large Economy with Aggregate Uncertainty

## Abstract

We study a large economy model in which individuals have private information about their productive abilities and their preferences. Moreover, there is aggregate uncertainty so that the social benefits from taxation and public goods provision are a priori unknown. The analysis is based on a mechanism design approach that imposes a requirement of robustness with respect to individual beliefs and a requirement of coalition-proofness. The paper provides a tractable and intuitive characterization of incentive-feasible tax and expenditure policies: Incentive constraints associated with productive abilities reflect only individual behavior, whereas those associated with public goods preferences reflect only collective behavior.

JEL Code: D71, D82, H21, H41.

Keywords: optimal taxation, public goods provision, revelation of preferences, robust mechanism design.

*Felix Bierbrauer*  
*Max Planck Institute, Bonn*  
*Kurt-Schumacher-Str. 10*  
*53113 Bonn*  
*Germany*  
*bierbrauer@coll.mpg.de*

May 18, 2009

I am very grateful for comments by Martin Hellwig, John Weymark and two anonymous referees. I also benefited from conversations with Ernesto Crivelli, Christoph Engel, Mike Golosov, Marco Sahm, Thomas Mertens, and Oleg Tsyvinski. I thank participants at the 2009 Decentralization Conference in St. Louis. Earlier versions of this paper circulated under the title “A unified approach to the revelation of public goods preferences and to optimal income taxation”. I am grateful for the hospitality of MIT where part of this research was conducted.

# 1 Introduction

This paper provides a characterization of tax and expenditure policies that can be implemented in a model in which tax revenues are used to cover the cost of public goods provision and in which individuals have private information about their preferences for public goods and about their productive abilities. It combines two different strands of the literature in public economics that have traditionally been separated: the theory of optimal income taxation in the tradition of Mirrlees (1971), and the theory of public goods provision under asymmetric information.

The theory of optimal income taxation is concerned with the characterization of a welfare-maximizing tax system under the assumption that individual labor supply is not exogenous but a function of the tax system itself. This literature mostly ignores public expenditures and assumes that the government has an exogenously given revenue requirement.<sup>1</sup>

The literature on public goods provision focusses on the characterization of a surplus-maximizing rule for the provision and financing of public goods under the assumption that individuals have private information on their preferences.<sup>2</sup> This literature, however, disregards the production side of the economy and the tax system as an alternative source of public goods finance.

A joint treatment of optimal taxation and preference revelation is difficult for two reasons. First, the combination of private information about productive abilities and of private information about public goods preferences gives rise to a multi-dimensional mechanism design problem.<sup>3</sup> Second, the models that are used in the theory of optimal taxation and the theory of public goods provision under asymmetric information are very different. While the former studies a large economy model in which each individual acts as a “price-taker” in the sense that the own behavior neither affects aggregate tax revenue nor public spending, the latter studies a finite economy in which each individual has a direct impact on the supply of public goods.

The contribution of this paper is to link these two different strands of the literatures. It is based on a large economy model with endogenous production, as is the theory of optimal taxation. The analysis of admissible tax and expenditure policies is complicated by the fact that there is aggregate uncertainty, i.e., the cross-sectional distribution of productive abilities and public goods preferences is taken to be random quantity. In particular, this implies that the social benefit from public goods provision is a priori unknown.

However, this problem of information aggregation may be considered trivial. In a large economy, a single individual’s valuation has no influence on the social benefit from public goods provision, so that a single individual’s communication of public goods preferences and abilities does not affect the supply of public goods. But if no individual can influence public goods

---

<sup>1</sup>The papers that include public goods assume that the social benefit from public goods provision is common knowledge, so that the problem of preference elicitation is moot; see, Boadway and Keen (1993), Gahvari (2006), Kaplow (2006) or Hellwig (2004).

<sup>2</sup>Early contributions ask whether efficiency is possible given private information on preferences, see Clarke (1971), Groves (1973) or d’Aspremont and Gérard-Varet (1979). The more recent literature includes participation constraints, e.g., Mailath and Postlewaite (1990), Hellwig (2003) or Neeman (2004).

<sup>3</sup>The literature also refers to such problems as multi-dimensional screening problems; see Rochet and Choné (1998) and Armstrong and Rochet (1999) for a theoretical treatment and Cremer et al. (2001) and Beaudry et al. (2009) for applications to the theory of optimal taxation.

provision anyway, then individuals may as well tell the truth. According to this view, to determine the social benefits from public goods provision is not a problem.

With an appeal to real-world political decision making, this conclusion seems to be overly optimistic. If one thinks about the role of political parties and special interest groups, the assumption that individuals with common interests may try to induce policies that are favorable to them seems more plausible than the alternative view that the problem of preference revelation is trivial provided that the number of individuals is sufficiently large. To articulate this concern in a model of optimal income taxation, this paper relies on a notion of coalition-proofness, which is based on the assumption that like-minded individuals can coordinate their communication of public goods preferences. If a sufficiently large number of individuals departs from truth-telling, the policy maker will end up with a wrong perception of the social benefits from public goods provision and the supply of public goods will be manipulated. A coalition-proof equilibrium requires that there is no group of individuals who can benefit from such a joint lie about their public goods preferences.

The main result of the paper is a tractable characterization of tax and expenditure policies that are implementable as a coalition-proof equilibrium, and, moreover, satisfy a robustness requirement with respect to the specification of individual beliefs (see Ledyard (1978) and Bergemann and Morris (2005)). The combination of coalition-proofness and robustness yields the following result: an implementable tax and expenditure policy must satisfy individual and collective incentive compatibility constraints. The individual incentive constraints ensure that individuals reveal their productive abilities. The collective incentive constraints ensure that individuals do not lie about their public goods preferences. These are one-dimensional incentive constraints in the following sense: For the characterization of individual incentive compatibility we may, without loss of generality, assume that individuals reveal their preferences. For collective incentive compatibility, we may assume that individuals reveal their productive abilities.

This decomposition into two sets of one-dimensional incentive constraints makes the mechanism design problem tractable. We illustrate this by means of a simple model of optimal income taxation and public goods provision, in which we characterize the optimal policy mechanism that satisfies both types of incentive constraints.

In addition, the paper shows that the relevance of the literature on the elicitation of public goods preferences is not limited to finite economy models in which each individual has a significant impact on the supply of public goods. Similar insights apply to large economy models in which individuals with aligned interests may jointly try to manipulate policy outcomes.

This paper contributes to a recent literature in public economics which uses a mechanism design approach in order to characterize optimal insurance contracts or tax systems.<sup>4</sup> The work that is most closely related to this paper is by Bassetto and Phelan (2008) and Kocherlakota and Phelan (2009). While there are differences with respect to the main applications, there is a shared methodological concern, namely how to deal with problems of equilibrium multiplicity

---

<sup>4</sup>See, for example, Golosov et al. (2003), Kocherlakota (2005), Acemoglu et al. (2008). Predecessors are Hammond (1979) and Guesnerie (1995).

in large economies with aggregate uncertainty.

A major difference, however, is the following: Bassetto and Phelan (2008) and Kocherlakota and Phelan (2009) are concerned with equilibrium multiplicity per se, while this paper is concerned with equilibrium multiplicity only to the extent that there is an alternative equilibrium which is attractive for a subset of individuals who are able to coordinate their behavior in such a way that the alternative equilibrium indeed becomes the more plausible outcome. The derivation of collective incentive compatibility constraints is based on the requirement that such alternative equilibria must not exist.

Moreover, the equilibrium multiplicity in Bassetto and Phelan (2008) is due to their focus on direct mechanisms. As is well known in the mechanism design literature, and also shown by Bassetto and Phelan (2008) for the setting of their paper, problems of equilibrium multiplicity that occur with direct mechanisms can often be overcome by enriching the set of mechanisms, i.e., by means of using non-direct mechanisms. For this reason, Kocherlakota and Phelan (2009) do not limit themselves to direct mechanisms and study a problem of equilibrium multiplicity that proves to be unavoidable even if more general mechanisms can be used.

In this paper, by contrast, equilibrium multiplicity can be addressed by restricting attention to direct mechanisms. Combining the requirements of robustness and coalition-proofness makes it possible to prove a revelation principle, so that it entails no loss of generality to look at tax and expenditure policies that can be implemented by a truthful coalition-proof equilibrium of a direct mechanism.

This result is important for the relevance of the individual and collective incentive compatibility constraints that are derived in this paper. They do not require specific assumptions about the “language” in which individuals communicate their characteristics to the mechanism designer. They might as well vote over public goods, instead of reporting their public goods preferences, and choose a utility-maximizing level of labor supply subject to an income tax schedule, instead of reporting their ability level. In any case, the outcome has to be individually and collectively incentive compatible.

The remainder of the paper is organized as follows. Section 2 specifies the model and contains a definition of robustness. In Section 3, we introduce the notion of a robust coalition-proof equilibrium and characterize implementable social choice functions. Section 4 studies implications of collective incentive compatibility requirements for more structured environments. The last section contains concluding remarks. All proofs are in the Appendix.

## 2 The Environment

There is a continuum of individuals. Individual  $i$ 's utility function is

$$U(q, c, y, w^i, \theta^i) = v(q, \theta^i) + u(z, w^i),$$

where  $q$  is the amount of a public good, and  $z = (c, y)$  is a bundle specifying an individual's private good consumption  $c$  and an individual's effective labor supply  $y$ .<sup>5</sup> The function  $u$  is increasing in  $c$ , decreasing in  $y$  and strictly quasiconcave. Moreover, it is assumed to be continuous. The function  $v$  gives the utility due to public good provision. For a given  $\theta^i$ ,  $v$  is twice continuously differentiable with  $v_q > 0$  and  $v_{qq} \leq 0$ .

The taste parameter  $\theta^i$  affects individual  $i$ 's valuation of the public good. We assume that there is a finite ordered set of possible taste parameters  $\Theta = \{\theta_1, \dots, \theta_m\}$  and that the marginal utility  $v_q$  from public good provision is increasing in the taste parameter,  $v_q(q, \theta_l) < v_q(q, \theta_{l+1})$ . Hence, individuals with a larger taste parameter benefit more from increased public good provision.

The skill or productivity parameter  $w^i$  affects individual  $i$ 's disutility of contributing to the economy's output. There is a finite ordered set of possible skill levels  $W = \{w_1, \dots, w_n\}$ . Indifference curves in a  $y - c$ -diagram satisfy the single crossing property, i.e., for any given  $y$  and  $c$ , and every  $k$ ,

$$-\frac{u_y(c, y, w_k)}{u_c(c, y, w_k)} > -\frac{u_y(c, y, w_{k+1})}{u_c(c, y, w_{k+1})}.$$

Less productive individuals need to exert more effort to produce a given output and hence need more compensation for an increased output requirement.

Individuals are privately informed about their taste parameter and about their skill level. Moreover, at an aggregate level there is uncertainty about the cross-sectional distribution of these characteristics in the economy. This distribution is taken to be the realization of a random variable which takes values in a set  $S$ . A typical element  $s \in S$  is referred to as a *state*. Formally,  $s$  is a matrix that contains the population shares of individuals who have certain characteristics, e.g.;  $s_{kl}$  is the population share of individuals with  $w^i = w_k$  and  $\theta^i = \theta_l$ .

The desirability of public good provision will typically depend on the state  $s$ . Consider two states  $s$  and  $s'$  such that, everything else being equal, in state  $s'$  more individuals have a high valuation of the public good than in state  $s$ . A welfare-maximizing policy will hence be such that in state  $s'$  more of the public good is provided than in state  $s$ .

A social choice function formalizes the dependence of outcomes on the state of the economy. It consists of a public goods provision rule  $q : s \mapsto q(s)$  that specifies for each state  $s$  how much of the public good is provided, and a private goods provision rule  $z : (s, w, \theta) \mapsto z(s, w, \theta)$ , which determines for each state, each skill level, and each taste parameter a consumption level and an output requirement, i.e.,  $z(s, w, \theta) = (c(s, w, \theta), y(s, w, \theta))$ .

A social choice function is said to be feasible, if for every  $s$ ,

$$\sum_{k=1}^n \sum_{l=1}^m (y(s, w_k, \theta_l) - c(s, w_k, \theta_l)) s_{kl} \geq r(q(s)), \quad (1)$$

where  $r$  is an increasing and convex cost function which captures the resource requirement of public good provision.

---

<sup>5</sup>The literature on the revelation of public goods preferences typically uses additively separable utility functions so that the utility from private goods consumption is a linear function of a good called money. Our model only assumes that the utility function is additively separable between the private goods  $c$  and  $y$ , on the one hand, and the public good  $q$ , on the other.

This specification restricts attention to social choice functions that are anonymous in the following sense: for a given state  $s$ , two individuals who have the same taste parameter and the same skill level receive the same private goods consumption and have to provide the same output. Also, a permutation of individual characteristics which leaves the joint cross-section distribution of public goods preferences and skill levels unaffected does neither affect the decision on public goods provision nor, for given  $(w, \theta)$ , the corresponding consumption level and output requirement.<sup>6</sup>

## 2.1 Robust Mechanism Design

We seek to characterize social choice functions that can be made the equilibrium allocation of a game of incomplete information in which individuals have information about their own characteristics but lack information about the characteristics of others and about the state  $s$  of the economy. Moreover, in order to avoid specific assumptions about individual beliefs, we will use the notion of robust implementability that has been introduced by Ledyard (1978) and Bergemann and Morris (2005).

An individual is henceforth characterized by a preference parameter, a productivity level, beliefs about the distribution of these characteristics in the economy and beliefs about the beliefs of other individuals. Following Bergemann and Morris (2005) we summarize these data in a *type space*  $\mathcal{T} := [T, \beta, \pi]$ , where  $T$  is the set of “abstract” types,  $\beta : T \rightarrow W \times \Theta$  is a mapping that specifies the skill level and the preference parameter of an individual with type  $t$ . We also refer to  $\beta(t)$  as the payoff type of type  $t$ . The function  $\pi$  specifies the beliefs of an individual of type  $t$  with respect to the cross-sectional distribution of types, i.e.,  $\pi(X | t)$  is the probability that a type  $t$  individual assigns to an event  $X \subset \Delta(T)$ .

In words, an individual’s type  $t^i$  is a random variable that takes values in the set of types  $T$ . Consequently, the payoff type and the belief type of individual  $i$  are given by the random variables  $\beta(t^i)$  and  $\pi(\cdot | t^i)$ , respectively. Also, the cross-sectional distribution of types  $\phi$  is a random variable that takes values in the set of probability measures over  $T$ ,  $\Delta(T)$ . An individual’s beliefs are specified as a probability distribution over the possible realizations of  $\phi$ .

In the following we treat the abstract set of types  $T$  and the function  $\beta$  as given. Also we assume that, for each element  $(w, \theta)$  of  $W \times \Theta$ , there exists a type such that  $\beta(t) = (w, \theta)$ . By contrast, we will vary the function  $\pi$  because we seek to understand how the outcomes that a mechanism designer can achieve depend on assumptions about the beliefs of individuals.

We seek to implement a social choice function by means of an anonymous allocation mechanism, i.e., a mechanism whose outcome is left unaffected by a permutation of individual actions.<sup>7</sup> Formally, an anonymous allocation mechanism  $M = (A, Q, Z)$  specifies an action set  $A$  and a collection of functions, where  $Q : \delta \mapsto Q(\delta)$  gives the level of public good provision as a function of the cross-sectional distribution of actions  $\delta$ , and the function  $Z : (\delta, a) \mapsto Z(\delta, a)$  specifies a consumption level  $C$  and an output requirement  $Y$  as a function of an individual’s message  $a$  and of the distribution of messages  $\delta$ ,  $Z(\delta, a) = (C(\delta, a), Y(\delta, a))$ .

<sup>6</sup>Guesnerie (1995) refers to these properties as recipient anonymity and anonymity in influence, respectively.

<sup>7</sup>For a more extensive discussion of anonymous games, see Kalai (2004).



A (mixed) strategy in the game induced by  $M$  is a function  $\sigma : T \rightarrow \Delta(A)$  that specifies a probability distribution over actions for each type of individual. Put differently, the action chosen by individual  $i$  is a random variable  $a(t^i)$ . The probability, conditional on the event  $t^i = t$ , that  $a(t^i)$  takes values in subset  $A'$  of  $A$  is in the following denoted by  $\sigma(A' | t)$ .

We denote by  $\delta(\phi, \sigma) := \phi \circ \sigma^{-1}$  the cross-sectional distribution of actions that is induced by  $\sigma$  if the cross-section distribution of types is  $\phi$ . We assume that a law of large numbers for large economies holds so that we can interpret  $\sigma(A' | t)$  both as the probability that the action chosen by a type  $t$  individual belongs to a subset  $A'$  of  $A$  and as the fraction of type  $t$  individuals who choose an action in  $A'$ .<sup>8</sup> Consequently, for given  $\phi$ , we can treat  $\delta(\phi, \sigma)$  as non-random.

Given that all other individuals follow a strategy  $\sigma$ , the expected payoff of a type  $t$  individual from randomizing over actions according to  $\chi \in \Delta(A)$  is denoted by

$$\tilde{U}(\sigma, \chi, t) := \int_{\Delta(T)} \int_A U(Q(\delta(\phi, \sigma)), Z(\delta(\phi, \sigma), a), \beta(t)) d\chi(a) d\pi(\phi | t) .$$

Given  $M$  and  $\pi$ , a strategy  $\sigma^*$  is said to be an interim Nash equilibrium, if for each  $t$ ,

$$\tilde{U}(\sigma^*, \sigma^*(t), t) \geq \tilde{U}(\sigma^*, \chi, t) , \tag{2}$$

for all  $\chi \in \Delta(A)$ .<sup>9</sup>

Given beliefs  $\pi$ , a social choice function  $(q, z)$  is said to be implementable if there is some mechanism  $M$  such that, for every  $\phi$ , the equilibrium allocation of the mechanism is equal to the outcome that is stipulated by the social choice function for the corresponding marginal distribution of payoff types, henceforth denoted by  $s(\phi) := \phi \circ \beta^{-1}$ . Formally, for given  $\pi$ , the mechanism  $M$  is said to implement the social choice function  $(q, z)$ , if  $\sigma^*$  is an interim Nash equilibrium, and moreover, for every  $\phi$ ,

$$Q(\delta(\phi, \sigma^*)) = q(s(\phi)) \tag{3}$$

and, for each  $i$ ,

$$Z(\delta(\phi, \sigma^*), a(t^i)) = z(s(\phi), \beta(t^i)) \tag{4}$$

with probability 1.

A social choice function is said to be *robustly implementable* if there is a mechanism  $M$  and a strategy  $\sigma^*$  such that conditions (2), (3), and (4) hold for every  $\pi$ .

This notion of robust implementability has been introduced by Ledyard (1978). It differs from the one used by Bergemann and Morris (2005), who allow the mechanism that implements the social choice function to depend on the beliefs  $\pi$ , i.e., a social choice function is robustly implementable in the sense of Bergemann and Morris (2005) if, for every  $\pi$ , there is a mechanism

---

<sup>8</sup>For a discussion of the law of large numbers in large economies, see Sun (2006), Al-Najjar (2004) or Judd (1985a).

<sup>9</sup>The terminology ‘‘interim Nash equilibrium’’ has been introduced by Bergemann and Morris (2005). A Bayes-Nash equilibrium is the special case that arises if individual beliefs  $\pi$  are derived from a common prior over the possible states of the economy.

$M(\pi)$  such that conditions (2), (3), and (4) hold. With our notion of robustness, the mechanism designer is not allowed to choose different mechanisms for different specification of beliefs.<sup>10</sup>

This difference can be illustrated if we focus on a special class of beliefs, namely those that lead to what is known in the literature as a mechanism design problem under complete information. For such problems it is assumed that individuals have complete information about their respective characteristics and that the mechanism designer is the only uninformed party.<sup>11</sup> Formally, we say that the function  $\pi$  gives rise to a complete information environment, if there exists a state  $s'$  such that, for all  $t$ ,  $\pi(\{\phi \mid s(\phi) = s'\} \mid t) = 1$ . The mechanism designer, by contrast, knows only that individuals know the cross-sectional distribution of payoff types. However, he does not know which one it is. Put differently, all he knows is that the function  $\pi$  belongs to the class  $\Pi^c$  of belief functions that generate a complete information environment. Consequently, it is natural to define *robust implementation under complete information* as the requirement that conditions (2), (3), and (4) hold for every  $\pi \in \Pi^c$ . In this environment, having a different mechanism for each  $\pi \in \Pi^c$  is not feasible, because the mechanism designer does not know  $\pi$ .

We can now turn to a characterization of the social choice functions that are robustly implementable.

**Proposition 1** *The following statements are equivalent.*

- (a) *A social choice function  $(q, z)$  is robustly implementable.*
- (b) *A social choice function  $(q, z)$  is robustly implementable under complete information.*
- (c) *A social choice function  $(q, z)$  satisfies the following individual incentive compatibility constraints: For every  $s \in S$  and every  $(w, \theta) \in W \times \Theta$ ,*

$$U(q(s), z(s, w, \theta), w, \theta) \geq U(q(s), z(s, \hat{w}, \hat{\theta}), w, \theta), \quad (5)$$

*for every  $(\hat{w}, \hat{\theta}) \in W \times \Theta$ .*

The proof of Proposition 1 is in the Appendix. It adapts arguments by Ledyard (1978) and Bergemann and Morris (2005) to the given large economy setup. The individual incentive compatibility constraints can be interpreted as follows: Suppose that the mechanism under consideration is a direct mechanism. Then, a truthful revelation of types must be an ex post equilibrium; i.e., once the state of the economy has been revealed, no individual regrets having reported his characteristics truthfully to the mechanism designer.

With the incentive constraints in (5), the problem of information aggregation – i.e., the problem to deduce the state  $s$  of the economy from the distribution of announcements that individuals make to the mechanism designer – is trivial. Given that the economy is large, a single

<sup>10</sup>With the solution concept of an interim Nash equilibrium, this distinction is not essential. Both versions of robustness give rise to the same constraints on the set of social choice functions, see Proposition 1 below. However, for the results in subsequent sections it makes a difference which notion of robustness is employed.

<sup>11</sup>See Moore (1992) for an overview.

individual cannot influence the state of the economy as perceived by the mechanism designer. As a consequence, there is no constraint that relates the outcomes prescribed for some state  $s$  to those prescribed for some alternative state  $s'$ . This is a striking difference to the literature on public goods provision in “small” economies in the tradition of Clarke (1971), Groves (1973) or d’Aspremont and Gérard-Varet (1979). In these models, each individual announces his public goods preferences and thereby also affects the cross-sectional distribution of announcements.

## 2.2 Optimal income taxation

Given Proposition 1, we focus in the following on individually incentive-compatible social choice functions. Using that the utility function  $U$  is additively separable, the incentive compatibility constraints in (5) can be equivalently written as follows: for every  $s \in S$  and every  $(w, \theta) \in W \times \Theta$ ,

$$u(c(s, w, \theta), y(s, w, \theta), w) \geq u(c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \hat{\theta}), w), \quad (6)$$

for all  $(\hat{w}, \hat{\theta})$ . The utility that individuals derive from public goods does not matter for incentive compatibility because (i) the economy is large and individuals take  $s$  as given, and (ii) by the separability assumption, an individual’s marginal rate of substitution between consumption  $c$  and output  $y$  does not depend on the supply of public goods. This also implies that differences in public goods preferences cannot be used as a screening device. Formally, the inequalities in (6) imply that, for every  $s$ , for every given  $w$  and every pair  $\theta$  and  $\theta'$ ,

$$u(c(s, w, \theta), y(s, w, \theta), w) = u(c(s, w, \theta'), y(s, w, \theta'), w), \quad (7)$$

so that two individuals who differ only in their taste parameter, derive the same utility from their respective  $(c, y)$  combination, in every state  $s$ .

Given condition (7), it is without loss of generality to assume that also  $c(s, w, \theta) = c(s, w, \theta')$  and  $y(s, w, \theta) = y(s, w, \theta')$ , for every  $s$ ,  $w$ , and every pair  $(\theta, \theta')$ .<sup>12</sup> In the following we may hence drop the dependence of consumption levels and output requirements on taste parameters and write simply  $c(s, w)$  and  $y(s, w)$ , respectively.

With this notation, we can write the individual incentive compatibility constraints as follows: for every  $s$ , every  $w$ , and every  $\hat{w}$ ,

$$u(c(s, w), y(s, w), w) \geq u(c(s, \hat{w}), y(s, \hat{w}), w). \quad (8)$$

The economy’s resource constraint in (1) can now be written as follows: For all  $s$ ,

$$\sum_{k=1}^n f_k(s) (y(s, w_k) - c(s, w_k)) \geq r(q(s)), \quad (9)$$

where, for any  $k \in \{1, \dots, n\}$ ,  $f_k(s) := \sum_{l=1}^m s_{kl}$ .

<sup>12</sup>Any welfare-maximizing social choice function is such that individual utility levels are generated at a minimal resource cost. Hence it must be true that  $y(s, w, \theta) - c(s, w, \theta) = y(s, w, \theta') - c(s, w, \theta')$ . This equality in conjunction with the fact that indifference curves in a  $y - c$  diagram are strictly increasing and strictly convex, yields  $c(s, w, \theta) = c(s, w, \theta')$  and  $y(s, w, \theta) = y(s, w, \theta')$ .

The model derived so far is a straightforward extension of the model of optimal income taxation by Mirrlees (1971).<sup>13</sup> A difference is that the present model has, in addition, private information on preferences and aggregate uncertainty so that the optimal policy depends on the state  $s$  of the economy. However, this has no bearing on the set of optimal policies. While the incentive compatibility constraints make it difficult to identify an individual's productivity level, information on individual preferences is irrelevant, and there is no constraint that complicates the determination of the state of the economy. The set of incentive-feasible policies is therefore equivalent to those in the theory of optimal income taxation.

### 3 Collective Incentive Compatibility

The characterization of individually incentive-compatible social choice functions in the preceding section has led to the conclusion that, in a large economy, uncertainty about the cross-sectional distribution of individual characteristics plays no role. For instance, if there is uncertainty about the social benefits from public good provision, this uncertainty can be easily resolved because, in a large economy, no individual has a direct influence on the supply of public goods, and hence no individual minds revealing his public goods preference.

This view is somewhat implausible. Every democratic society knows the constant struggle of political parties and all sorts of special interest groups about the desirability of expenditure policies. Hence, in a large economy, individuals undertake collective actions in order to influence the provision of public goods. This suggests that, in addition to *individual* incentive compatibility constraints which incorporate how individual behavior depends on policy choices, the model should also include *collective* incentive compatibility constraints which take account of the possibility that individuals may lie collectively about their preferences in order to induce policies that are more favorable to them.

Our approach is as follows: A coalition is a subset of types who have an incentive to coordinate the actions they choose. If they jointly deviate from the equilibrium strategy they may affect the mechanism designer's perception of the state of the economy and induce outcomes that are more favorable to them. In addition, we impose the requirement that individual behavior under such a coordinated deviation is a best response to the behavior of all other individuals. We say that a social choice function is collectively incentive-compatible if it can be implemented by a mechanism that does not provoke such deviations.

Before we develop these concepts more generally, we first use a simple model of optimal income taxation and public goods provision to illustrate that, if we take only individual incentive compatibility constraints into account, there may indeed be scope for manipulative collective actions.

---

<sup>13</sup>Contributions to the theory of optimal income taxation often use a primal approach which relies on a characterization of social choice functions that can be achieved by means of an income tax system, see Stiglitz (1982), Boadway and Keen (1993), Gahvari (2006), Hellwig (2007), or Bierbrauer (2009a). This yields implementability conditions that, for a given  $s$ , coincide with the constraints in (9) and (8).

### 3.1 A simple model of optimal income taxation and public good provision

Suppose that there are two states of the economy,  $s^1$  and  $s^2$ . In both states half of the population is half-skilled and half of the population is low-skilled. Moreover, in both states, among the low-skilled, fifty per cent have a high valuation of the public good. Among the high-skilled, however, 25 per cent have a high taste for the public good in state  $s^1$  and 75 per cent have a high taste for the public good in state  $s^2$ .

We assume that preferences take the following form,

$$U(q, c, y, \theta, w) = \theta q + \ln(c) - \frac{y}{w},$$

where  $q \in \{0, 1\}$  is the provision level of an indivisible public good. Moreover, we restrict attention to affine income tax schedules, that consist of a linear income tax rate  $\tau$  and a lump sum transfer  $\alpha$ .<sup>14</sup>

A mechanism specifies, for each state  $s \in \{s^1, s^2\}$ , whether the public good is provided,  $q(s)$ , the income tax rate,  $\tau(s)$ , and the lump sum transfer,  $\alpha(s)$ . Given the tax rate  $\tau(s)$  and the lumps sum transfer  $\alpha(s)$  individuals with skill level  $w$  choose  $c$  and  $y$  in order to maximize  $\ln(c) - \frac{y}{w}$  subject to the budget constraint  $c = (1 - \tau(s))y + \alpha(s)$ . Consequently, for given  $s$  and  $w$ ,  $c(s, w)$  and  $y(s, w)$  are given by

$$\{c(s, w), y(s, w)\} = \operatorname{argmax}_{c, y} \ln(c) - \frac{y}{w} \text{ s.t. } c = (1 - \tau(s))y + \alpha(s) \quad (10)$$

By a standard revealed preferences argument, the allocation that results from these utility maximization problems is individually incentive-compatible; i.e., it satisfies the inequalities in (8), for every  $s$ . The feasibility constraint takes the following form: for each  $s$ ,

$$\tau(s) \left( \frac{1}{2}y(s, w_1) + \frac{1}{2}y(s, w_2) \right) \geq kq(s) + \alpha(s), \quad (11)$$

where  $k$  is the per capita cost of public goods provision.

Suppose that, for every state  $s$ , we choose  $q(s)$ ,  $\tau(s)$  and  $\alpha(s)$  in order to maximize utilitarian welfare

$$W(s) := \frac{1}{2} \left( \ln(c(s, w_1)) - \frac{y(s, w_1)}{w_1} \right) + \frac{1}{2} \left( \ln(c(s, w_2)) - \frac{y(s, w_2)}{w_2} \right)$$

subject to the feasibility constraint in (11) and the constraints in (10), which ensure that individual behavior is utility maximizing given  $\tau(s)$  and  $\alpha(s)$ . It is a straightforward exercise to verify that the solution is such that:

1. For any  $\tau(s)$  and  $\alpha(s)$ ,  $c(s, w_k) = w_k(1 - \tau(s))$  and  $y(s, w_k) = w_k - \frac{\alpha(s)}{1 - \tau(s)}$ .
2. The optimal tax rate  $\tau^*(s)$  and transfer  $\alpha^*(s)$  are such that

$$\tau^*(s) = \tau^* = 1 - (\bar{w}\lambda)^{-1} \quad \text{and} \quad \alpha^*(s) = (1 - \tau^*) (\bar{w}\tau^* - kq(s)),$$

---

<sup>14</sup>This is a simple version of the model of optimal linear income taxation that has been introduced by Sheshinski (1972). The results below can also be obtained in a model with a more general information structure, a more general specification of preferences and unrestricted non-linear income taxation, see Bierbrauer (2009b).

where  $\bar{w} := \frac{1}{2}(w_1 + w_2)$  is the average skill level and  $\lambda := \frac{1}{2}(\frac{1}{w_1} + \frac{1}{w_2})$  is the average marginal disutility of an increased output requirement. Note that switching from  $q(s) = 0$  to  $q(s) = 1$  implies that the lump sum transfer goes down, whereas the tax rate  $\tau$  remains constant.

We assume in the following that public goods provision is desirable in state  $s^2$  where many individuals have a high valuation of the public good, but not in state  $s^1$  where only few individuals have a high valuation of the public good.<sup>15</sup> In the following, we argue that there may be an incentive for the group of high-skilled individuals to lie about their public goods preferences.

High-skilled individuals benefit, *ceteris paribus*, more from public goods provision than low-skilled individuals. The reason is that the high-skilled individuals do not suffer as much from the cost of public goods provision. To see this formally, note that as we switch from non-provision to provision of the public good, the individuals' consumption remains constant, whereas the term  $\frac{y}{w_k}$  increases by  $\frac{k}{w_k}$ . Obviously, this expression is decreasing in the skill level. Consequently, it may well be the case that high-skilled individuals benefit from public goods provision even if their valuation of the public good is low. This is the case if

$$\theta_1 > \frac{k}{w_2}. \quad (12)$$

Suppose that the mechanism designer uses a direct mechanism to learn what the state of the economy is; i.e., each individual reports a taste parameter and a skill level to the mechanism designer. If the inequality in (12) holds, then all high-skilled individuals are better off if the mechanism designer chooses policy under the assumption that many high-skilled individuals have a high valuation of the public good. Moreover, high-skilled individuals can induce this outcome if they behave as follows: they communicate their skill parameters truthfully but, irrespective of their true taste parameter, declare a low taste parameter with probability  $\frac{1}{4}$  and a high taste parameter with probability  $\frac{3}{4}$ . This implies that the messages received by the mechanism designer are as if the true state of the economy was  $s^2$ .

### 3.2 Coalition-Proof Equilibrium and the Revelation Principle

We now define formally what it means that the game induced by a mechanism  $M = (A, Q, Z)$  has a coalition-proof equilibrium for some given specification of beliefs  $\pi$ .<sup>16</sup>

Given  $M$  and  $\pi$ , a strategy  $\sigma^*$  is said to be a coalition-proof interim Nash equilibrium if it is an interim Nash equilibrium and there exists no other interim Nash equilibrium  $\sigma'$  with the following properties: There exists a subset of types  $R \subset T$  such that the following requirements are fulfilled:

<sup>15</sup>This is the case provided that  $\frac{1}{4}\theta_1 + \frac{3}{4}\theta_2 > \lambda k > \frac{3}{4}\theta_1 + \frac{1}{4}\theta_2$ .

<sup>16</sup>The definition below is a somewhat simplified version of the notion of a coalition-proof Nash equilibrium that by Bernheim et al. (1986). In particular, we also take a non-cooperative approach to coalition formation and require that the individual behavior that is prescribed by a deviation is individually a best response. For reasons of tractability, however, we do not require that deviations are relevant only in so far as they do not trigger deviations by subcoalitions, which themselves are a concern only if they do not provoke deviations by subcoalitions of the initial subcoalition, etc.

- i) Given that all types in  $T \setminus R$  play according to  $\sigma^*$ , it is a best response for a type in  $R$  to play according to  $\sigma'$ . Formally, for all  $t \in R$ , and all  $\chi \in \Delta(A)$ ,

$$\tilde{U}((\sigma_{T \setminus R}^*, \sigma'_R), \sigma'(t), t) \geq \tilde{U}((\sigma_{T \setminus R}^*, \sigma'_R), \chi, t), \quad (13)$$

where  $\sigma_{T \setminus R}^*$  is the restriction of  $\sigma^*$  to types not in  $R$ , and  $\sigma'_R$  is the restriction of  $\sigma'$  to types in  $R$ .

- ii) The outcome that is induced if all types in  $T \setminus R$  play according to  $\sigma^*$ , and all types in  $R$  play according to  $\sigma'$ , is preferred by all individuals with types in  $R$ ; i.e., for all  $t \in R$ ,

$$\tilde{U}((\sigma_{T \setminus R}^*, \sigma'_R), \sigma'(t), t) > \tilde{U}(\sigma^*, \sigma^*(t), t). \quad (14)$$

- iii) The outcome that is induced if all types play according to  $\sigma'$  is preferred by all individuals with types in  $R$ . Formally, for all  $t \in R$ ,

$$\tilde{U}(\sigma', \sigma'(t), t) > \tilde{U}(\sigma^*, \sigma^*(t), t). \quad (15)$$

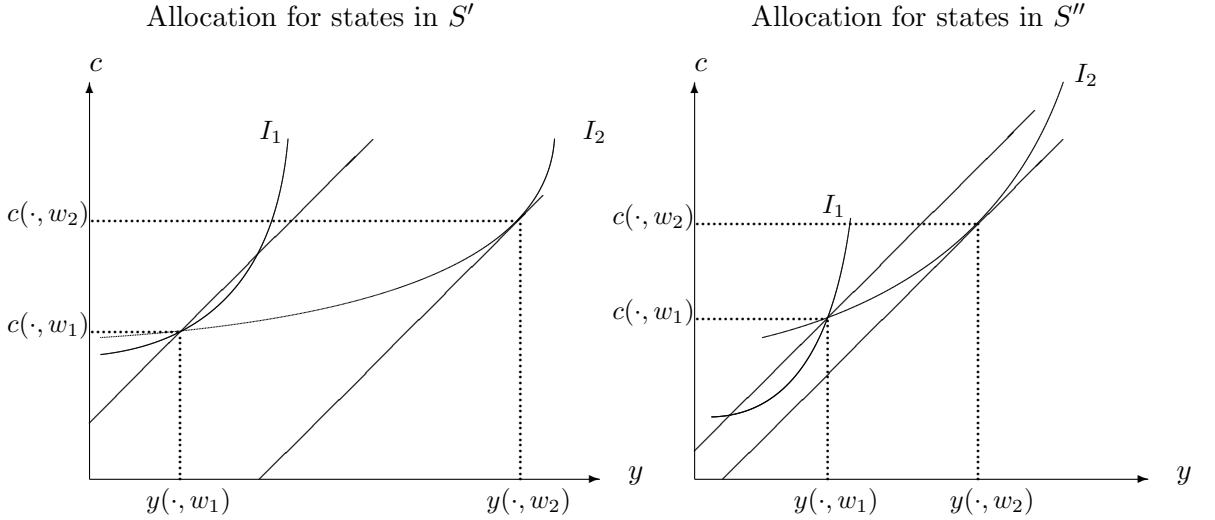
The underlying idea is that an equilibrium  $\sigma^*$  must not leave incentives for a subset of individuals to coordinate their behavior in such a way that they induce an outcome that makes all of them better off. The above definition is very demanding with respect to the consistency requirements that such a deviation from an equilibrium strategy  $\sigma^*$  has to satisfy: The behavior that is prescribed by the deviation must be a best response for the deviating types, both if all types play according to  $\sigma'$  and also if only the types in  $R$  play according to  $\sigma'$ . Also, in both cases, the outcome that is induced by the deviation must be beneficial for types in  $R$ . Hence, individuals with types in  $R$  are willing to behave according to  $\sigma'$  and are thereby made better off, provided that all types in  $R$  play according to  $\sigma'$ . This does not depend on whether or not types in  $T \setminus R$  also switch to  $\sigma'$  or stay with the old equilibrium  $\sigma^*$ . From the perspective of types in  $R$ , getting a preferred outcome is therefore a pure coordination problem.

Given  $\pi$ , a social choice function  $(q, z)$  is said to be implementable as a coalition-proof interim Nash equilibrium, if there is a mechanism  $M$  and a strategy  $\sigma^*$  such that (i)  $\sigma^*$  is a coalition-proof interim Nash equilibrium, and (ii) the equilibrium allocation coincides with the prescription of the social choice function for every  $\phi$ ; i.e., conditions (3) and (4) are satisfied. We say that a social choice function is robustly implementable if there is a mechanism  $M$  and a strategy  $\sigma^*$  such that requirements (i) and (ii) are fulfilled for every  $\pi$ .

**Proposition 2** *If a social choice function is robustly implementable as a coalition-proof interim Nash equilibrium, then there is also a direct mechanism (i.e., a mechanism with  $A = T$ ) such that truth-telling (i.e., the “honest” strategy  $h : T \rightarrow \Delta(T)$  with  $h(\{t\} | t) = 1$ , for all  $t$ ) robustly implements  $(q, z)$  as a coalition-proof interim Nash equilibrium.*

The Proposition establishes that the revelation principle holds. This result is of interest for the following reason: It is known in the literature that that the objective to implement a social choice

Figure 1: The social choice function



function as the unique equilibrium of some mechanism typically requires the use of non-direct mechanisms.<sup>17</sup> Hence, according to this literature, if one worries about equilibrium multiplicity, considering only direct mechanisms involves a loss of generality. Proposition 2, by contrast, says that if, under a direct mechanism, there is an alternative equilibrium so that some individuals have incentives to lie, then there is no mechanism at all that implements the given social choice function in a coalition-proof way. The difference is due to the robustness requirement that is imposed in this paper. We illustrate this by means of the following example. A formal proof of the Proposition is in the Appendix.

**Example.** Suppose that there are no public goods. Moreover, there are only two possible skill levels,  $W = \{w_1, w_2\}$ . Let  $S'$  be the set of states in which the population share of the high-skilled is larger or equal than  $\frac{1}{2}$ , and let  $S''$  be the set of states in which it is smaller than  $\frac{1}{2}$ . Suppose that we seek to implement a social choice function with the following properties: for all states, there is a binding incentive compatibility constraint so that high-skilled individuals are indifferent between the bundles  $z(s, w_1) = (c(s, w_1), y(s, w_1))$  and  $z(s, w_2) = (c(s, w_2), y(s, w_2))$ , and there is redistribution from the high-skilled to the low-skilled,  $y(s, w_2) - c(s, w_2) > 0$  and  $y(s, w_1) - c(s, w_1) < 0$ .<sup>18</sup> However, the level of redistribution varies across states; it is large for all  $s \in S'$  and small for all  $s \in S''$ . This is illustrated by Figure 1. In this Figure,  $I_1$  is the relevant indifference curve of the low-skilled, and  $I_2$  is the one of the high-skilled individuals.

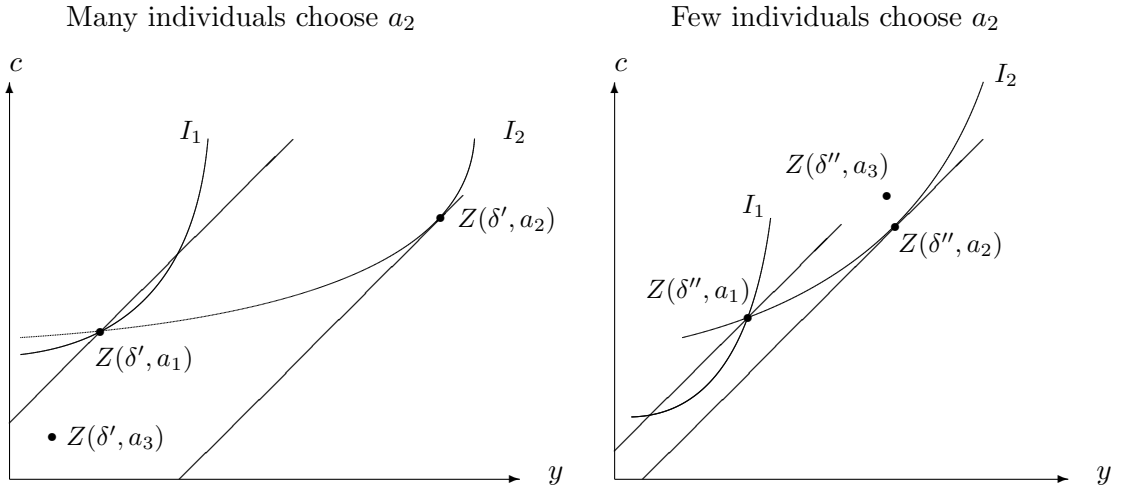
With a direct mechanism, this social choice function cannot be implemented in a coalition-proof way. Since the states in  $S'$  involve more redistribution than the states in  $S''$ , we have that  $u(z(s', w_2), w_2) < u(z(s'', w_2), w_2)$ , for any pair of states  $s' \in S'$ , and  $s'' \in S''$ . Hence, a

<sup>17</sup>For a recent application, see Bassetto and Phelan (2008). For a more general discussion, see Jackson (2001).

<sup>18</sup>As is well-known in the literature, an optimal utilitarian allocation typically has these properties, see, e.g., Hellwig (2007).



Figure 2: The non-direct mechanism



collective lie so that some of the high-skilled individuals declare a low-skill level, makes all high-skilled individuals better off. Moreover, for every  $s$ , the incentive constraint of the high-skilled is binding, so that the high-skilled are giving a best response if they lie about their skill level.

In the following we will show that a mechanism designer who knows the function  $\pi$  can generally eliminate the scope for such a collective deviation. We will then argue that this is no longer true if we insist on robustness.

For the sake of concreteness, suppose first that all high-skilled individuals believe that the states in  $S'$  and  $S''$  are equally likely, and, in addition, that this is known by the mechanism designer. He may then use a non-direct mechanism  $M = (A, Z)$ , with  $A = \{a_1, a_2, a_3\}$  which works as follows: In any state, an individual who chooses action  $a_1$  receives the consumption-income bundle dedicated to the low-skilled individuals; i.e., taking the action  $a_1$  is interpreted as saying “I am a low-skilled individual”. Likewise, an individual who chooses action  $a_2$  receives the bundle intended for the high-skilled individuals. The action  $a_3$  gives a very unattractive bundle if many individuals choose action  $a_2$ . This is illustrated in the left part of Figure 2, where  $\delta'$  indicates a distribution of actions so that more than half of the population chooses  $a_2$ . However, if only few individuals choose action  $a_2$ , then action  $a_3$  is very attractive for the high-skilled individuals. This is illustrated by the right part of Figure 2, where  $\delta''$  indicates a distribution of actions so that less than half of the population chooses  $a_2$ .

This non-direct mechanism implements the social choice function in a coalition-proof way, because there is no longer an equilibrium in which high-skilled individuals communicate having a low-skill level: If all high-skilled individuals anticipate that the number of individuals choosing  $a_2$  is small, so that the payoffs associated with actions  $a_1$ ,  $a_2$ , and  $a_3$  are as in the right part of Figure 2, then they are individually better off by choosing action  $a_3$ . By contrast, if many individuals choose  $a_2$ , then payoffs are as in the left part, and  $a_2$  is a best response for the high-skilled individuals.

This construction eliminates the equilibrium in which high-skilled individuals mimic low-skilled individuals. Conditional on many high-skilled individuals choosing action  $a_1$  it is a best response for a high-skilled individual to choose action  $a_3$ . By contrast, the other equilibrium in which all high-skilled individuals choose action  $a_2$  and thereby communicate their skill level truthfully to the mechanism designer remains intact: If the situations in Figure 2 arise with equal probability, then choosing  $a_2$  is a best response for the high-skilled because action  $a_3$  yields a very bad outcome with probability  $\frac{1}{2}$ . Consequently, the non-direct mechanism  $M = (A, Z)$  implements the social choice function under consideration in a coalition-proof way.

However, the non-direct mechanism does no longer work if we insist on robustness. To see this, suppose that all individuals believe that states in  $S''$  occur with probability 1. Then, all individuals believe that the payoffs associated with actions  $a_1$ ,  $a_2$ , and  $a_3$  are as in the right part of Figure 2. Consequently, all individuals with a high-skill level will choose action  $a_3$  instead of action  $a_2$ , so that the social choice function is no longer reached.

### 3.3 Partial Coalition-Proofness

We will show in the following that the requirement of coalition-proofness can be considerably simplified. More specifically, we show that it suffices to address incentives for groups of individuals to misrepresent their public goods preferences. There is no need to worry about groups of individuals who might lie about their skill levels.

We need to introduce some notation and some definitions. First consider the following modification of the incentive compatibility constraints in (8): for every  $s$ , every  $w$ , and every  $\hat{w}$ ,

$$u(z(s, w), w) \geq u(z(s, \hat{w}), w) + \epsilon, \quad (16)$$

for some given  $\epsilon > 0$ . These constraints require that, for every  $s$ , an individual with skill level  $w$  prefers the “own” consumption-output bundle  $z(s, w)$  strictly over any alternative bundle  $z(s, \hat{w})$  where the parameter  $\epsilon$  is the minimal intensity of this strict preference.

Also, for direct mechanisms, we define a *partially* coalition-proof interim Nash equilibrium as follows: There is no deviation from truth-telling, which satisfies the consistency requirements (13)-(15), so that the deviating individuals reveal their skills but lie about their taste parameters. We refer to such a deviation henceforth as a partial lie. Formally, a partial lie is a function  $l_p : R \rightarrow \Delta(T)$  which is such that, for every  $t \in R$ ,  $l_p(\{t' \mid w(t') = w(t)\} \mid t) = 1$ . Also, we denote by  $\hat{\phi}(\phi, l_p)$  the announced distribution of types that is induced by  $l_p$  if the true distribution of types is  $\phi$  and the types not in  $R$  follow the truthful strategy  $h$ .

Finally, we denote the set of social choice functions which satisfy the modified incentive compatibility constraints in (16), the feasibility constraints in (1), and, moreover, are robustly implementable as a partially coalition-proof interim Nash equilibrium, by  $\Omega(\epsilon)$ .

**Proposition 3** *Suppose that there is some  $\bar{\epsilon} > 0$  so that the set  $\bigcup_{0 < \epsilon \leq \bar{\epsilon}} \Omega(\epsilon)$  is compact. Then, to every social choice function  $(q, z)$  that is robustly implementable as a partially coalition-proof interim Nash equilibrium, there is a social choice function  $(q', z')$  that is robustly implementable as a coalition-proof interim Nash equilibrium and is such that for every  $s$  and every  $(\theta, w)$ ,  $v(q(s), \theta) + u(z(s, w), w)$  is arbitrarily close to  $v(q'(s), \theta) + u(z'(s, w), w)$ .*

Partial coalition-proofness is a weaker requirement than “full” coalition-proofness. The former requirement excludes lies about taste parameters, whereas the latter requirement excludes lies about taste parameters and skill levels. The Proposition says that, one may, nevertheless focus on weaker requirement of partial coalition-proofness. The reason is as follows: the adapted incentive constraints in (16) imply that a lie about skill levels is incompatible with individual behavior being a best response. From an ex post perspective, individuals would always regret a lie about their skill level. Hence, the constraints in (16) in conjunction with partial coalition-proofness imply full coalition-proofness. Moreover, if we let  $\epsilon$  become very small, we can approximate any social choice function in  $\Omega(0)$ . This shows that the set of social choice functions that are partially coalition-proof and the set of social choice functions that are fully coalition proof are “approximatively” equivalent. The following example shows why there is no “exact” equivalence.

**The Example continued.** We have shown that the social choice function that is illustrated in Figure 1 cannot be implemented as a robust coalition-proof equilibrium because the high-skilled have an incentive to lie. We will now argue that there is, however, a social choice function which is arbitrarily close and does not face this problem.

Suppose that the social choice function in Figure 1 is modified as follows: In both graphs, the bundle for high-skilled individuals is moved to a slightly higher indifference curve.<sup>19</sup> This implies that truth-telling is the unique best response of the high-skilled, for every state  $s$ . A deviation that involves lies about skill levels is therefore no longer consistent with equilibrium behavior.

The example illustrates the general insight in Proposition 3. Once we introduce a tiny amount of slack into the incentive compatibility constraints, deviations that involve lies about skill levels are no longer viable. The example also shows why the slack is needed. If incentive compatibility constraints are binding, lies that involve skill levels are a concern.

### 3.4 The main result

We will turn to the main result of the paper, which establishes that the constraints on the set of implementable social choice functions can be separated as follows: The requirement of robust implementability as an interim Nash equilibrium yields *individual* incentive compatibility constraints. These constraints ensure that individuals communicate their skills truthfully to the mechanism designer. The requirement of coalition-proofness yields *collective* incentive compatibility constraints, which ensure that individuals reveal their public goods preferences. In particular, there is no need to worry about individual incentives for a misrepresentation of public goods preferences or collective incentives for a misrepresentation of skill levels. As we will see in Section 3.5, this makes the multi-dimensional mechanism design problem tractable.

**Proposition 4** *An individually incentive compatible social choice function is robustly implementable as a partially coalition-proof equilibrium if and only if the following collective incentive-*

---

<sup>19</sup>To preserve feasibility, we may simultaneously have to move the low-skilled individuals to a slightly lower indifference curve.

compatibility constraint holds for every  $\pi$ : There is no subset of types  $R \subset T$  with a partial lie  $l_p : R \rightarrow \Delta(T)$  so that for all  $t \in R$ ,

$$\begin{aligned} & \int_{\Delta(T)} U(q(s(\hat{\phi}(\phi, l_p))), z(s(\hat{\phi}(\phi, l_p))), \beta(t), \beta(t)) d\pi(\phi | t) \\ & > \int_{\Delta(T)} U(q(s(\phi)), z(s(\phi)), \beta(t), \beta(t)) d\pi(\phi | t). \end{aligned}$$

The proof of the Proposition is straightforward. We have shown before that individual incentive compatibility implies that individuals who differ only in their public goods preferences must receive the same bundle of private goods, see equation (7). Consequently, individuals are willing to communicate any taste parameter to the mechanism, so that every partial lie leads to a new equilibrium. This implies that partial coalition-proofness reduces to the requirement that there is no subset of individuals who benefit from a partial lie.

Consequently, the requirement of coalition-proofness can be reduced to a simple set of inequality constraints: There must not exist a group of individuals who benefit from the change of the allocation that is induced by a false communication of public goods preferences. This simple characterization is available because, by Proposition 3, as far as coalition-proofness is concerned, we may, without loss of generality, assume that productive abilities are communicated truthfully. Put differently, the problem to determine the state  $s$  of the economy requires appropriately calibrated incentives only in so far as there is uncertainty about the cross-sectional distribution of public goods preferences.

Individual incentive compatibility, by contrast, deals with a pure screening problem; i.e., with the revelation of individuals' characteristics, once the state  $s$  of the economy has been determined. As we have seen in Section 2, this requires appropriate incentives only for the revelation of productive abilities but not for public goods preferences.

### 3.5 Optimal income taxation and public good provision subject to collective incentive compatibility constraints

To illustrate the tractability of our framework, we will now revisit the model of optimal income taxation and public good provision that was introduced in Section 3.1, and which gives rise to the following collective incentive problem: High-skilled individuals exaggerate their public goods preferences because the public good is crowding out redistribution, and high-skilled individuals benefit from decreased redistribution even if they do not value the public good very much.

The role of the collective incentive constraints is to eliminate the incentives for such behavior. In the present setting, these constraints take the following form,

$$\theta_1 + \ln(c(s^2, w_2)) - \frac{y(s^2, w_2)}{w_2} \leq \ln(c(s^1, w_2)) - \frac{y(s^1, w_2)}{w_2}. \quad (17)$$

The constraint says that high-skilled individuals with a low taste for the public good weakly prefer the allocation for state  $s^1$  in which the public good is not provided over the allocation in state  $s^2$  in which it is provided.

The following Proposition characterizes the social choice function that maximizes expected utilitarian welfare (assuming for simplicity that the objective function gives equal weights to

both states),  $EW = \frac{1}{2}W(s^1) + \frac{1}{2}W(s^2)$ , subject to the feasibility constraints in (11), the implementability constraints in (10), and the collective incentive compatibility constraint in (17).

**Proposition 5** *Suppose the optimum has  $q(s^1) = 0$  and  $q(s^2) = 1$ . Then  $\tau(s^1)$  and  $\alpha(s^1)$  are lower and  $\tau(s^2)$  and  $\alpha(s^2)$  are higher as compared to the case without collective incentive constraints.*

A formal proof of the Proposition can be found in the Appendix. The intuitive explanation is as follows: Collective incentive compatibility constraints are needed to ensure that the high-skilled individuals do not have an incentive to exaggerate their public goods preferences. Put differently, from the perspective of the high-skilled, the allocation involving public good provision is too attractive relative the one involving non-provision. Collective incentive compatibility therefore requires that the state with public good provision is made more attractive and the state with non-provision is made less attractive. This is achieved via an adjustment of the level of redistribution. Public good provision is accompanied by increased redistribution and higher taxes and non-provision by less redistribution and lower taxes; that is, the collective incentive compatibility constraints induce a complementarity between public good provision and the level of redistribution.

## 4 Collective incentive compatibility under complete information

In this last section, we study the implications of the collective incentive compatibility constraints for mechanism design problems under conditions of complete information. This assumption facilitates a comparison to the literature on public goods that uses models with finitely many individuals. This comparison will demonstrate that the relevance of this literature is not limited to economies which are so small that each single individual has a noticeable impact on the supply of public goods. The same ideas are applicable in a large economy in which individuals with aligned interests can jointly manipulate policy outcomes.

We assume that there is a commonly known cross-sectional distribution of productive abilities, i.e., the vector  $f = (f_1, \dots, f_n)$  is treated as known. By Proposition 3 this is without loss of generality. Aggregate uncertainty is therefore entirely due to the fact that the distribution of public goods preferences among individuals with skill level  $w_k$ , is, for each  $k \in \{1, \dots, n\}$ , unknown. Formally, the list  $s_k := (s_{k1}, \dots, s_{km})$  is an unknown quantity, which takes values in a set  $S_k$ . In the following we will occasionally write  $s = (s_{-k}, s_k) \in (S_{-k}, S_k)$  in order to emphasize the impact of individuals with skill level  $w_k$  on the (true or perceived) state of the economy.

We also assume that the environment gives rise to a mechanism design problem under complete information; i.e., the state  $s$  of the economy is assumed to be known among individuals, but not to the mechanism designer. The mechanism designer only knows that  $\pi \in \Pi^c$ .

**Proposition 6** *If a social choice function is robustly implementable as a coalition-proof interim*

*Nash equilibrium, this implies that it is also robustly implementable under complete information. The converse implication does not hold.*

By Proposition 1, if we only require robust implementability as a Nash equilibrium, then, it does not make a difference whether or not we limit attention to complete information environments. Proposition 6 shows that, if we add the requirement of coalition-proofness, the equivalence between robust implementability and robust implementability under complete information is lost.

The reason is that, with specific beliefs, joint manipulations by individuals who have antagonistic preferences may restrict the set of implementable social choice functions. For instance, a joint manipulation by individuals who are harmed by public good provision and by individuals who benefit from public good provision may work as follows. Those who are harmed by public good provision are willing to lie so as to make public good provision possible in some state  $s'$  if they are compensated by non-provision in some other state  $s''$ . Likewise, individuals who benefit from public good provision are willing to block public good provision in some state if this implies that the public good will be provided in some other state. However, this is not possible under conditions of complete information. For instance, if all individuals put a lot of probability mass on the true state of the economy being  $s'$ , then individuals who oppose public good provision, will not consider the above manipulation attractive. Likewise, if individuals assume that the state is almost surely  $s''$ , then individuals who benefit from public good provision will not participate. Consequently, if attention is limited to complete information environments, then collective incentive compatibility is a less stringent constraint.

#### 4.1 Representative Individuals

In the following, we focus on a specific set of collective incentive compatibility constraints, namely those which address lies by individuals who have the same skill level.<sup>20</sup> We will provide conditions so that the corresponding collective incentive compatibility constraints can be interpreted as referring to a representative individual with skill level  $w_k$  who has to be given incentives so as to ensure that he reports his own preference truthfully. The collective incentive compatibility constraints therefore admit an interpretation as individual incentive constraints for an economy with  $n$  agents who have private information about their preferences.

To prevent partial lies by individuals who all have skill level  $w_k$ , the following collective incentive compatibility constraints have to be satisfied: For any  $s_{-k}$  and any pair  $s_k$  and  $s'_k$ , there must be  $l$  with  $s_{kl} > s'_{kl}$  so that

$$v(q(s_{-k}, s_k), \theta_l) + V_k(s_{-k}, s_k) \geq v(q(s_{-k}, s'_k), \theta_l) + V_k(s_{-k}, s'_k), \quad (18)$$

where  $V_k(s) := u(z(s, w_k), w_k)$  is a shorthand notation for the utility from private goods that individuals with skill level  $w_k$  realize in state  $s$ . If this constraint was violated then, in state  $s = (s_{-k}, s_k)$ , all individuals with a taste parameter  $\theta_l$  so that  $s_{kl} > s'_{kl}$  would benefit from a

---

<sup>20</sup>These conditions are necessary but not sufficient for collective incentive compatibility, because the possibility of partial lies by individuals of different abilities is left open.

lie that makes the mechanism designer believe that the distribution of public goods preferences among individuals with skill level  $w_k$  equals  $s'_k$ .

The following assumption implies that the collective incentive compatibility constraints in (18) admit the convenient characterization via a representative individual.

**Assumption 1** *Suppose that, for each skill level  $w_k$ , the set  $S_k$  contains  $m$  elements that are ordered according to first order stochastic dominance;  $S_k = \{s_k^1, \dots, s_k^m\}$  where, for any  $l$ ,  $s_k^{l+1}$  first order stochastically dominates  $s_k^l$ . Moreover, as one moves from  $s_k^l$  to  $s_k^{l+1}$  the mass of all taste parameters other than  $\theta_l$  and  $\theta_{l+1}$  is unaffected; i.e., the mass of individuals with taste parameter  $\theta_{l+1}$  is increased at the expense of individuals with taste parameter  $\theta_l$ .*

**Proposition 7** *Under Assumption 1, a social choice function satisfies the constraints in (18) if and only if for all  $k$ , for all  $s_{-k}$  and for all  $l$ ,*

$$v(q(s_{-k}, s_k^l), \theta_l) + V_k(s_{-k}, s_k^l) \geq v(q(s_{-k}, s_k^{l+1}), \theta_l) + V_k(s_{-k}, s_k^{l+1}), \quad (19)$$

and

$$v(q(s_{-k}, s_k^l), \theta_l) + V_k(s_{-k}, s_k^l) \geq v(q(s_{-k}, s_k^{l-1}), \theta_l) + V_k(s_{-k}, s_k^{l-1}). \quad (20)$$

Condition (19) is a local upward incentive compatibility constraint for individuals with skill level  $w_k$  and taste parameter  $\theta_l$ . If  $s_k^l$  is the true distribution of preferences among those with skill level  $w_k$ , then a subset of these individuals could report a taste parameter equal to  $\theta_{l+1}$  so that the distribution of preferences that is communicated to the mechanism designer equals  $s_k^{l+1}$ . The incentive constraint (19) ensures that this is not attractive for these individuals. Similarly, (20) is a local downward incentive compatibility constraint that ensures that these individuals do not understate their preferences. The Proposition shows that these local incentive constraints are sufficient to ensure that the constraints in (18) are satisfied.<sup>21</sup>

The incentive conditions in (20) and (19) can be interpreted as referring to a “representative agent” in the following sense: consider a special version of the above model, so that  $s_k = s_k^1$  indicates that all individuals with  $w^i = w_k$  have a taste parameter equal to  $\theta_1$ ,  $s_k = s_k^2$  indicates that all individuals with  $w^i = w_k$  have a taste parameter equal to  $\theta_2$ , etc. Hence, in every state of the economy, all individuals with the same skill level also have the same taste parameter. Now assume that in order to determine  $s_k$  the policy maker asks a randomly chosen individual with skill level  $w_k$  to make an announcement about the own taste parameter. This again gives rise to the incentive conditions (20) and (19). These constraints can therefore be interpreted as standard dominant strategy incentive compatibility constraints in a finite model with  $n$  representative agents, one for each skill level, who make an announcement about their own taste parameter.

---

<sup>21</sup>This follows from standard arguments which exploit the assumption that the marginal utility from public good provision  $v_q$  increases in the taste parameter, see the Appendix.

## Implications for the analysis of income tax systems

We now use Proposition 7 to derive two implications of collective incentive compatibility for the joint analysis of tax and expenditure policies. These are necessary conditions, i.e., if they are violated public policy will be manipulated.

By (19) and (20), collective incentive compatibility holds if and only if for all  $k$ , for all  $s_{-k}$  and for all  $l < m$ ,

$$v_q(q(s_k^{l+1}, s_{-k}), \theta_{l+1}) - v_q(q(s_k^l, s_{-k}), \theta_{l+1}) \geq V_k(s_k^l, s_{-k}) - V_k(s_k^{l+1}, s_{-k}), \quad (21)$$

and

$$V_k(s_k^l, s_{-k}) - V_k(s_k^{l+1}, s_{-k}) \geq v_q(q(s_k^{l+1}, s_{-k}), \theta_l) - v_q(q(s_k^l, s_{-k}), \theta_l). \quad (22)$$

According to (21), the drop in the utility from private goods  $V_k$  as one moves from  $s_k^l$  to  $s_k^{l+1}$  must be bounded from above so that individuals with taste parameter  $\theta_{l+1}$  are willing to accept this utility loss in exchange for the utility gain that is due to increased public good provision. If this condition was violated then individuals with taste parameter  $\theta_{l+1}$  would understate their preferences for the public good.

According to (22), the drop in the utility from private goods must be bounded from below. If this property was violated then the cost of increased public good provision that is born by individuals with taste parameter  $\theta_l$  is so small that these individuals would exaggerate their public goods preferences.

Adding these inequalities and using the assumption that the marginal utility from public good provision increases in the taste parameter reveals that  $q(s_k^l, s_{-k}) \leq q(s_k^{l+1}, s_{-k})$ , i.e., if there are ceteris paribus more individuals with a high taste parameter among those with skill level  $w_k$ , this increases the amount of the public good that is provided. Using (22) this implies that  $V_k(s_k^l, s_{-k}) \geq V_k(s_k^{l+1}, s_{-k})$ , if more of the public good is provided this must imply that individuals with skills  $w_k$  pay more taxes and hence derive a lower utility from private goods. Moreover, (21) and (22) imply that

$$q(s_k^l, s_{-k}) = q(s_k^{l+1}, s_{-k}) \iff V_k(s_k^l, s_{-k}) = V_k(s_k^{l+1}, s_{-k}).$$

The implication  $q(s_k^l, s_{-k}) = q(s_k^{l+1}, s_{-k}) \implies V_k(s_k^l, s_{-k}) = V_k(s_k^{l+1}, s_{-k})$  can be interpreted as “no taxation without representation”, i.e., if a larger willingness to pay for the public good among individuals with skills  $w_k$  is not reflected by the provision rule  $q$ , then it is not possible to tax these individuals more heavily. They would never admit a higher valuation of the public good if the only consequence was that they have to pay more taxes. The converse implication reads as “no representation without taxation.” If individuals do not pay more taxes as their valuation of the public good goes up, then the provision rule for the public good cannot incorporate this change in preferences. Otherwise, individuals would always exaggerate their preferences for the public good since they do not have to pay for it.

## 4.2 Groves Mechanisms

In the following we relate the collective incentive compatibility conditions to in (20) and (19) to the analysis of Groves mechanisms. In finite economy models, these are mechanisms which



ensure that a truthful revelation of public goods preferences is a dominant strategy equilibrium and that public goods are provided in a surplus-maximizing way.

In the present setting, the surplus from providing  $x$  units of the public good in state  $s$  is given by  $\sum_{k=1}^n f_k \sum_{l=1}^m s_{kl} v(x, \theta_l) - r(x)$ . The surplus maximizing provision rule is in the following denoted by  $q^*(s)$ . A Groves mechanism is such that, for each  $k$  and  $s_k$ ,

$$V_k(s_k^j, s_{-k}) = \sum_{g \neq k} v(q^*(s_k^j, s_{-k}), \theta(s_g)) - nr(q^*(s_k^j, s_{-k})) .$$

One well-known problem of Groves mechanism is that budget balance is typically not possible.<sup>22</sup> Moreover, surplus-maximizing public goods provision is typically not desirable if there are equity concerns, as is typically the case in a Mirrleesian model of optimal income taxation. That being said, the analysis of Groves mechanisms has proven to be very instructive for the analysis of revelation mechanisms. This lends some interest to Proposition 8 which states necessary and sufficient conditions so that, in a large economy model in which individuals may jointly hide their true public goods preferences, a Groves mechanism may simultaneously induces truth-telling and achieve a surplus maximizing public goods supply.

**Proposition 8** *A Groves mechanism implements provision rule  $q^*$  subject to the collective incentive compatibility constraints (20) and (19) if and only if the following assumptions are met: Skills are uniformly distributed and, conditional on a certain skill level, there is no heterogeneity in public goods preferences.*

The if-part of this Proposition is not surprising. If all representative agents receive the same “weight” ( $f_k = \frac{1}{n}$ , for all  $k$ ), and moreover there is no heterogeneity in public goods preferences among the represented individuals, then we have a model that is equivalent to a finite economy model with  $n$  individuals. For these models, it is well understood that a Groves mechanism works. The only if-part, however, shows that this structure is not only sufficient but also necessary. If we want to implement the surplus maximizing provision rule, then a Groves mechanism is, generally, incompatible with the incentive compatibility constraints in (20) and (19).

## 5 Concluding Remarks

This paper has analyzed a large economy in which individuals are privately informed about their productive abilities and their preferences for public goods. Moreover, there is aggregate uncertainty with respect to the cross-sectional distribution of these characteristics. Using a mechanism design approach that involves a requirement of robustness with respect to assumptions on individual beliefs as well as a requirement of coalition-proofness, the analysis has identified two sets of incentive conditions for public policy. Individual incentive compatibility constraints take into account how individuals respond to an income tax system that determines their after-tax income as a function of their labor supply. Collective incentive compatibility constraints take

<sup>22</sup>Green and Laffont (1977) show that this is generally true for any mechanism with a dominant strategy equilibrium that achieves surplus maximization.

care of the possibility that individuals may lobby for certain tax and expenditure policies and thus addresses the political reactions that may be triggered by the policy mechanism.

The analysis has led to a set of necessary conditions that are implied by the requirement of collective incentive compatibility. If a group of individuals experiences a shift in their public goods preferences such that their willingness to pay for a public good is increased, then it must be true that more of the public good is provided (otherwise these individuals understate their public goods preferences) and that these individuals pay more taxes (otherwise they exaggerate their preferences). More generally speaking, the tax system confronts individuals with prices for public goods. Collective incentive compatibility requires that these prices are set in an “appropriate” manner, namely in such a way that the demand for public goods can be determined.

While such arguments are familiar from mechanism design approaches to the free-rider problem in public goods provision, they have not been introduced into the literature on optimal taxation. This paper’s contribution is to develop a framework that makes it possible to address problems of preference elicitation and optimal taxation simultaneously.

The argument in this paper is based on a model of optimal income taxation. However, a similar reasoning applies to any model of taxation. To illustrate this, imagine a model of capital taxation in the sense of Chamley (1986) and Judd (1985b). These are dynamic models in which a policy maker faces a choice between a tax on labor income and tax on capital income. Chamley and Judd show that, in a steady state equilibrium, there should be no tax on capital income. However, the arguments of this paper suggest, that if people with capital income do not contribute to the financing of public goods, then it becomes impossible to reflect their preferences when choosing the level of public expenditures (“no representation without taxation”). Alternatively, if their preferences are to be considered, then people with capital income will have to pay taxes at least in those circumstances where they value the public good highly. Hence, it seems that a tax on capital income will be needed if it is desirable that people with capital income have an influence on the provision of public goods.

## References

- Acemoglu, D., Golosov, M., and Tsyvinski, A. (2008). Markets versus governments. *Journal of Monetary Economics*, 55:159–189.
- Al-Najjar, N. (2004). Aggregation and the law of large numbers in large economies. *Games and Economic Behavior*, 47:1–35.
- Armstrong, A. and Rochet, J. (1999). Multi-dimensional screening: A user’s guide. *European Economic Review*, 43:959–979.
- Bassetto, M. and Phelan, C. (2008). Tax riots. *Review of Economic Studies*, 75:649–669.
- Beaudry, P., Blackorby, C., and Szalay, D. (2009). Taxes and employment subsidies in optimal redistribution programs. *American Economic Review*, 99:216–242.
- 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. (2009a). A note on optimal income taxation, public goods provision and robust mechanism design. *Journal of Public Economics*, 93:667–670.
- Bierbrauer, F. (2009b). Optimal income taxation and public good provision with endogenous interest groups. *Journal of Public Economic Theory*, 11:311–342.
- Boadway, R. and Keen, M. (1993). Public goods, self-selection and optimal income taxation. *International Economic Review*, 34:463–478.
- Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica*, 54:607–622.
- Clarke, E. (1971). Multipart pricing of public goods. *Public Choice*, 11:17–33.
- Cremer, H., Pestieau, P., and Rochet, J. (2001). Direct versus indirect taxation: The design of the tax structure revisited. *International Economic Review*, 42:781–799.
- d’Aspremont, C. and Gérard-Varet, L. (1979). Incentives and incomplete information. *Journal of Public Economics*, 11:25–45.
- Gahvari, F. (2006). On the marginal costs of public funds and the optimal provision of public goods. *Journal of Public Economics*, 90:1251–1262.
- Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). Optimal indirect and capital taxation. *Review of Economic Studies*, 70:569–587.
- Green, J. and Laffont, J. (1977). Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica*, 45:472–487.
- 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. (2007). A contribution to the theory of optimal utilitarian income taxation. *Journal of Public Economics*, 91:1449–1477.

- Jackson, M. (2001). A crash course in implementation theory. *Social Choice and Welfare*, 18:655–708.
- Judd, K. (1985a). The law of large numbers with a continuum of i.i.d. random variables. *Journal of Economic Theory*, 35:19–25.
- Judd, K. (1985b). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28:59–83.
- Kalai, E. (2004). Large robust games. *Econometrica*, 72:1631–1665.
- Kaplow, L. (2006). Public goods and the distribution of income. *European Economic Review*, 50:1627–1660.
- Kocherlakota, N. (2005). Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation. *Econometrica*, 73:1587–1621.
- Kocherlakota, N. and Phelan, C. (2009). On the robustness of laissez-faire. *Journal of Economic Theory*, forthcoming.
- Ledyard, J. (1978). Incentive compatibility and incomplete information. *Journal of Economic Theory*, 18:171–189.
- Mailath, G. and Postlewaite, A. (1990). Asymmetric bargaining procedures with many agents. *Review of Economic Studies*, 57:351–367.
- 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.
- Neeman, Z. (2004). The relevance of private information in mechanism design. *Journal of Economic Theory*, 117:55–77.
- Rochet, J. and Choné, P. (1998). Ironing, sweeping, and multidimensional screening. *Econometrica*, 66:783–826.
- Sheshinski, E. (1972). The optimal linear income-tax. *Review of Economic Studies*, 39:297–302.
- Stiglitz, J. (1982). Self-selection and Pareto-efficient taxation. *Journal of Public Economics*, 17:213–240.
- Sun, Y. (2006). The exact law of large numbers via fubini extension and characterization of insurable risks. *Journal of Economic Theory*, 126:31–69.

# A Appendix

## Proof of Proposition 1

By the revelation principle, a social choice function  $(q, z)$  is implementable as an interim Nash equilibrium by some mechanism  $M$  for given  $\pi$ , if and only if it is truthfully implementable, i.e., if and only if there exists a direct mechanism  $M'$  with an action set  $A' = T$  and outcome functions  $Q' : \phi \mapsto Q'(\phi)$ , and  $Z' : (\phi, t) \mapsto Y'(\phi, t)$  such that (i) Truth-telling is an interim Nash equilibrium; i.e., for all  $t$

$$t \in \operatorname{argmax}_{t' \in T} \int_{\Delta(T)} U(Q'(\phi), Z'(\phi, t'), \beta(t)) d\pi(\phi | t); \quad (23)$$

and (ii) the equilibrium allocation is equal to the social choice function,

$$Q'(\phi) = q(s(\phi)) \quad (24)$$

and, for every  $t$ ,

$$Z'(\phi, t) = z(s(\phi), \beta(t)). \quad (25)$$

Consequently, we may say that a social choice function is robustly implementable (robustly implementable under complete information) if and only if there is a direct mechanism satisfying (23), (24) and (25) for every  $\pi$  (for every  $\pi \in \Pi^c$ ).

We first show that  $(c) \Rightarrow (a)$ . Given an incentive compatible social choice function  $(q, z)$  and given an arbitrary  $\pi$ , construct a direct mechanism  $(T, Q', Z')$  such that (24) and (25) hold. We seek to verify that, for every  $t$ ,

$$\begin{aligned} t &\in \operatorname{argmax}_{t' \in T} \int_{\Delta(T)} U(Q'(\phi), Z'(\phi, t'), \beta(t)) d\pi(\phi | t) \\ &= \operatorname{argmax}_{t' \in T} \sum_s \hat{\pi}(s | t) U(q(s), z(s, \beta(t')), \beta(t)), \end{aligned}$$

where  $\hat{\pi}(s' | t) := \pi(\{\phi \in \Delta(T) | s(\phi) = s'\} | t)$ . Equivalently, for every  $t$ ,

$$\beta(t) \in \operatorname{argmax}_{(w', \theta') \in W \times \Theta} \sum_s \hat{\pi}(s | t) U(q(s), z(s, w', \theta'), \beta(t)).$$

This follows from the fact that  $(q, z)$  is incentive compatible.

The implication  $(a) \Rightarrow (b)$  is trivial. To complete the proof it remains to be shown that  $(b) \Rightarrow (c)$ . Fix  $\pi \in \Pi^c$  and suppose that a direct mechanism  $(T, Q', Z')$  truthfully implements  $(q, z)$ . Let  $s'$  be the corresponding cross-sectional distribution of payoff types. Using conditions (24) and (25) to substitute for  $Q'$ , and  $Z'$ , the equilibrium conditions in (23) becomes: for all  $t$  and all  $t'$ ,

$$U(q(s'), z(s', \beta(t)), \beta(t)) \geq U(q(s'), z(s', \beta(t')), \beta(t));$$

or, equivalently, for all  $(w, \theta)$  and  $(w', \theta')$ ,

$$U(q(s), z(s, w, \theta), w, \theta) \geq U(q(s), z(s, w', \theta'), w, \theta).$$

Since the choice of  $s'$  was arbitrary, the latter inequality holds for all  $s \in S$ . Hence,  $(q, z)$  is individually incentive compatible. ■

## Proof of Proposition 2

The Proposition follows from Lemmas 1 - 5 below.

**Lemma 1** *Consider a mechanism  $M = (A, Q, Z)$  and a strategy  $\sigma^*$  that robustly implements a social choice function  $(q, z)$  as a coalition-proof interim Nash equilibrium. Let  $A(t, \sigma^*)$  be the set of actions that are chosen by type  $t$  with positive probability under  $\sigma^*$ . Then, for every  $\phi$ , and every  $t$ ,*

$$A(t, \sigma^*) = \operatorname{argmax}_{a \in A} U(Q(\delta(\phi, \sigma^*)), Z(\delta(\phi, \sigma^*), a), \beta(t)) . \quad (26)$$

**Proof** A necessary condition for robust implementability as a coalition-proof interim Nash equilibrium, is robust implementability as an interim Nash equilibrium. In particular, this implies that the strategy  $\sigma^*$  prescribes behavior that is a best response for every  $\pi$  that puts unit mass on some fixed but arbitrary  $\phi$ . If the Lemma is false, then there exists at least one such  $\phi$  so that the behavior of at least one type is not a best response to the behavior of all others. Hence, a contradiction. ■

Given Lemma 1 we may, without loss of generality, limit attention to mechanisms that satisfy condition (26).

The proof the revelation principle proceeds as follows: We fix a social choice function  $(q, z)$  and suppose that there is a mechanism  $M = (A, Q, Z)$  and a strategy  $\sigma^*$  that satisfy condition (26) as well as conditions (3) and (4). From the revelation principle for the solution concept of “robust implementability as an interim Nash equilibrium”, we know that there is also a direct mechanism  $M' = (T, Q', Z')$  such that the “honest” strategy  $h$  satisfies all these properties. We then fix an arbitrary  $\pi$  and show that  $\sigma^*$  is coalition-proof in the game induced by  $M$  only if  $h$  is coalition-proof in the game induced by  $M'$ .

The proof is by contradiction. Hence, suppose to the contrary that, for some given  $\pi$ ,  $\sigma^*$  is coalition-proof, but  $h$  is not. Since  $M$  and  $\sigma^*$  as well as  $M'$  and  $h$  satisfy (3) and (4) we have that, for each  $\phi$ ,

$$Q(\delta(\phi, \sigma^*)) = Q'(\phi) = q(s(\phi)) , \quad (27)$$

and, for each  $t$ ,

$$Z(\delta(\phi, \sigma^*), a(t^i)) = Z'(\phi, t^i) = z(s(\phi), \beta(t^i)) , \quad (28)$$

with probability 1.

Also, since  $h$  fails to be coalition-proof there exists a subset of types  $R$  and a non-truthful equilibrium strategy (a “lie”)  $l : T \rightarrow \Delta(T)$  such that, for all  $t \in R$ , the consistency requirements (13) – (15) are satisfied.

The lie  $l$  induces, for each  $\phi \in \Delta(T)$ , an announced cross-sectional distribution of types  $\hat{\phi}(\phi)$  with

$$\hat{\phi}(T' | \phi) = \int_{t \in T} l(T' | t) d\phi(t)$$

for any subset  $T'$  of  $T$ .

With reference to  $l$  we now define a strategy  $\sigma'$  for the game induced by mechanism  $M$  in the following way: For every  $t$  and every subset  $A'$  of  $A$ , let

$$\sigma'(A' | t) = \int_{t' \in T} \sigma^*(A' | t') dl(t' | t) . \quad (29)$$

This construction ensures that the distribution of actions that results if all individuals behave according to  $\sigma'$  and the distribution of types is given by  $\phi$  equals the distribution of actions that is induced

if all individuals behave according to  $\sigma^*$  and the distribution of types equals  $\hat{\phi}(\cdot | \phi)$ . Formally, for every  $\phi$ ,

$$\delta(\phi, \sigma') = \delta(\hat{\phi}(\phi), \sigma^*) . \quad (30)$$

To see this, note that for every subset  $A'$  of  $A$ ,

$$\begin{aligned} \delta(A' | \hat{\phi}(\phi), \sigma^*) &= \int_{t' \in T} \sigma^*(A' | t') d\hat{\phi}(t' | \phi) \\ &= \int_{t' \in T} \sigma^*(A' | t') d \left( \int_{t \in T} l(t' | t) d\phi(t) \right) \\ &= \int_{t \in T} \int_{t' \in T} \sigma^*(A' | t') dl(t' | t) d\phi(t) \\ &= \int_{t \in T} \sigma'(A' | t) d\phi(t) \\ &= \delta(A' | \phi, \sigma') . \end{aligned}$$

**Lemma 2** *For the given  $\pi$ ,  $\sigma'$  is an interim Nash equilibrium for the game induced by  $M$ .*

**Proof** Suppose otherwise. Then there is a type  $t$  and an action  $a \in A$  such that

$$\begin{aligned} &\int_{\Delta(T)} U(Q(\delta(\phi, \sigma')), Z(\delta(\phi, \sigma'), a), \beta(t)) d\pi(\phi | t) \\ &> \int_{\Delta(T)} U(Q(\delta(\phi, \sigma')), Z(\delta(\phi, \sigma'), \hat{a}), \beta(t)) d\pi(\phi | t) , \end{aligned}$$

for all  $\hat{a} \in A(t, \sigma')$ . By (30) we may equivalently say that there is a type  $t$  and an action  $a \in A$  such that

$$\begin{aligned} &\int_{\Delta(T)} U(Q(\delta(\hat{\phi}(\phi), \sigma^*)), Z(\delta(\hat{\phi}(\phi), \sigma^*), a), \beta(t)) d\pi(\phi | t) \\ &> \int_{\Delta(T)} U(Q(\delta(\hat{\phi}(\phi), \sigma^*)), Z(\delta(\hat{\phi}(\phi), \sigma^*), \hat{a}), \beta(t)) d\pi(\phi | t) , \end{aligned} \quad (31)$$

for all  $\hat{a} \in A(t, \sigma')$ . By Lemma 1, we may, without loss of generality, assume that the action  $a$  that enters the left hand side of (31) belongs to  $A(t, \sigma^*)$ . (Note that if (31) holds for some  $a$ , then it must hold, in particular, for all  $a \in A(t, \sigma^*)$ .) Also, by construction of  $\sigma'$ ,  $\hat{a} \in A(t, \sigma')$  only if there is a subset  $T'$  of  $T$  with  $l(T' | t) > 0$  and  $\hat{a} \in A(t', \sigma^*)$  for all types  $t' \in T'$ . Using these observations in conjunction with equations (27) and (28) implies that there is  $t' \in T'$  so that

$$\begin{aligned} &\int_{\Delta(T)} U(Q'(\hat{\phi}(\phi)), Z'(\hat{\phi}(\phi), t), \beta(t)) d\pi(\phi | t) \\ &> \int_{\Delta(T)} U(Q'(\hat{\phi}(\phi)), Z'(\hat{\phi}(\phi), t'), \beta(t)) d\pi(\phi | t) . \end{aligned}$$

This contradicts the assumption that  $l$  is an interim Nash equilibrium for mechanism  $M'$ . ■

It remains to be verified that the equilibrium strategy  $\sigma'$  satisfies the consistency requirements (13)-(15).

**Lemma 3** *For the given  $\pi$ , the equilibrium  $\sigma'$  under mechanism  $M$  is payoff equivalent to the equilibrium  $l$  under mechanism  $M'$ ; i.e., for each  $t$ ,*

$$\begin{aligned} &\int_{\Delta(T)} \int_A U(Q(\delta(\phi, \sigma')), Z(\delta(\phi, \sigma'), a), \beta(t)) d\sigma'(a | t) d\pi(\phi | t) \\ &= \int_{\Delta(T)} \int_{t' \in T} U(Q'(\hat{\phi}(\phi)), Z'(\hat{\phi}(\phi), t'), \beta(t)) dl(t' | t) d\pi(\phi | t) . \end{aligned} \quad (32)$$

**Proof** Using equations (29) and (30) the left-hand-side of equation (32) can be equivalently written as

$$\begin{aligned} & \int_{\Delta(T)} \int_A U(Q(\delta(\hat{\phi}(\phi), \sigma^*)), Z(\delta(\hat{\phi}(\phi), \sigma^*), a), \beta(t)) d \left( \int_{t' \in T} \sigma^*(a | t') dl(t' | t) \right) d\pi(\phi | t) \\ &= \int_{\Delta(T)} \int_{t' \in T} \int_A U(Q(\delta(\hat{\phi}(\phi), \sigma^*)), Z(\delta(\hat{\phi}(\phi), \sigma^*), a), \beta(t)) d\sigma^*(a | t') dl(t' | t) d\pi(\phi | t). \end{aligned}$$

Moreover, equations (27) and (28) imply that

$$\begin{aligned} & \int_A U(Q(\delta(\hat{\phi}(\phi), \sigma^*)), Z(\delta(\hat{\phi}(\phi), \sigma^*), a), \beta(t)) d\sigma^*(a | t') \\ &= U(Q'(\hat{\phi}(\phi)), C'(\hat{\phi}(\phi), t'), Y'(\hat{\phi}(\phi), t'), \beta(t)), \end{aligned}$$

with probability 1. This establishes the Lemma. ■

Lemma 3, in conjunction with the fact that equilibrium  $\sigma^*$  under  $M$  is payoff equivalent to equilibrium  $h$  under  $M'$ , implies that  $\sigma'$  satisfies the consistency requirement (15).

To establish that also the consistency requirements (13) and (14) are fulfilled, we assume in the following that only types in  $R$  behave according to  $\sigma'$ , whereas types in  $T \setminus R$  stick to  $\sigma^*$ .

For each  $\phi$  this induces an announced cross-sectional distribution of types  $\bar{\phi}(\phi)$  with

$$\bar{\phi}(T' | \phi) = \int_{t \in R} l(T' | t) d\phi(t) + \int_{t \in T \setminus R} h(T' | t) d\phi(t)$$

for any subset  $T'$  of  $T$ .

Also, we have that, for every  $\phi$ , the distribution of actions that results if individuals with types in  $R$  behave according to  $\sigma'$  and all other follow  $\sigma^*$  equals the distribution of actions that results if all individuals follow  $\sigma^*$  and the distribution of types equals  $\bar{\phi}(\phi)$ . Formally, for every  $\phi$ ,

$$\delta(\bar{\phi}(\phi), \sigma^*) = \delta(\phi, (\sigma_{T \setminus R}^*, \sigma'_R)). \quad (33)$$

To see that this true, note that for any subset  $A'$  of  $A$ ,

$$\begin{aligned} \delta(A' | \bar{\phi}(\phi), \sigma^*) &= \int_{t' \in T} \sigma^*(A' | t') d\bar{\phi}(t | \phi) \\ &= \int_{\{t' \in T\}} \sigma^*(A' | t') d \left( \int_{t \in R} l(t' | t) d\phi(t) + \int_{t \in T \setminus R} h(t' | t) d\phi(t) \right) \\ &= \int_{t \in R} \int_{t' \in T} \sigma^*(A' | t') dl(t' | t) d\phi(t) + \int_{t \in T \setminus R} \int_{t' \in T} \sigma^*(A' | t') dh(t' | t) d\phi(t) \\ &= \int_{t \in R} \sigma'(A' | t) d\phi(t) + \int_{t \in T \setminus R} \sigma^*(A' | t) d\phi(t) \\ &= \delta(A' | \phi, (\sigma_{T \setminus R}^*, \sigma'_R)). \end{aligned}$$

**Lemma 4** *For the given  $\pi$  and the given mechanism  $M$ , it is a best response for types in  $R$  to follow  $\sigma'$  if types not in  $R$  follow  $\sigma^*$ .*

Lemma 4 implies that the consistency requirement (13) is satisfied. A formal proof is omitted, since it requires only a straightforward adaptation of the proof of Lemma 2. (One just has to replace  $\sigma'$  by  $(\sigma_{T \setminus R}^*, \sigma'_R)$  and  $\hat{\phi}(\phi)$  by  $\bar{\phi}(\phi)$ , and evaluate the inequalities for some type  $t \in R$ .)

To complete the proof of Proposition 2 we need to show that also consistency requirement (14) is satisfied. This is established by the following Lemma. Again we omit a formal proof, because it follows the same logic as the proof of Lemma 3.



**Lemma 5** For the given  $\pi$ , and any  $t \in R$ , the outcome of strategy profile  $(\sigma_{T \setminus R}^*, \sigma'_R)$  under mechanism  $M$  is payoff equivalent to the outcome of strategy profile  $(h_{T \setminus R}, l_R)$  under  $M'$ .

### Proof of Proposition 3

Consider a social choice function  $(q, z) \in \Omega(0)$ . Since  $\bigcup_{0 \leq \epsilon \leq \bar{\epsilon}} \Omega(\epsilon)$  is compact, we can construct a sequence of social choice functions  $\{(q^k, z^k)\}_{k=1}^\infty$  with  $(q^k, z^k) \in \Omega(\frac{\epsilon}{k})$ , for each  $k$ , which converges to  $(q, z)$ . By continuity of the functions  $v$  and  $u$  this implies that also, for each  $s, w$ , and  $\theta$ ,  $v(q^k(s), \theta) + u(z^k(s, w), w)$  converges to  $v(q(s), \theta) + u(z(s, w), w)$ .

To complete the proof we show that  $(q^k, z^k) \in \Omega(\epsilon)$  for  $\epsilon > 0$  implies that  $(q^k, z^k)$  is not only partially but also fully coalition-proof. Suppose otherwise, then there must be an alternative equilibrium in which some individuals lie about their skills. However, by the modified incentive constraints in (16) revealing the true taste parameter is a strictly dominant strategy. Hence, a contradiction. ■

### Proof of Proposition 4

Since the social choice is individually incentive-compatible, it satisfies condition (7). As a consequence, any partial lie satisfies the consistency requirements (13). Also, it is an equilibrium if types in  $R$  behave according to a partial lie and types not in  $R$  behave truthfully. The consistency requirement (14) therefore implies that the consistency requirement (15) is satisfied. We can therefore conclude that robust partial coalition-proofness holds if and only if the consistency requirement (14) fails for all  $R, l_p$  and  $\pi$ . This is the statement of the Proposition. ■

### Proof of Proposition 5

The implementability constraints in (10) imply that for any  $\tau(s)$  and  $\alpha(s)$ ,  $c(s, w_k) = w_k(1 - \tau(s))$  and  $y(s, w_k) = w_k - \frac{\alpha(s)}{1 - \tau(s)}$ . Substituting these expressions into the constraints and the objective function, and dropping constant terms that are irrelevant for the maximization problem, yields the following optimization problem: Choose  $\tau(s^1)$ ,  $\alpha(s^1)$ ,  $\tau(s^2)$ , and  $\alpha(s^2)$  in order to maximize

$$\ln(1 - \tau(s^1)) + \lambda \frac{\alpha(s^1)}{1 - \tau(s^1)} + \ln(1 - \tau(s^2)) + \lambda \frac{\alpha(s^2)}{1 - \tau(s^2)}$$

subject to the feasibility constraints,

$$\bar{w}\tau(s) = \frac{\alpha(s)}{1 - \tau(s)} + kq(s),$$

for each  $s$ , and the collective incentive compatibility constraint,

$$\theta_1 + \ln(1 - \tau(s^2)) + \frac{1}{w_2} \frac{\alpha(s^2)}{1 - \tau(s^2)} \leq \ln(1 - \tau(s^1)) + \frac{1}{w_2} \frac{\alpha(s^1)}{1 - \tau(s^1)}.$$

Upon using the feasibility constraints to substitute for  $\frac{\alpha(s^1)}{1 - \tau(s^1)}$  and  $\frac{\alpha(s^2)}{1 - \tau(s^2)}$  this optimization problem can be rewritten as: Choose  $\tau(s^1)$  and  $\tau(s^2)$  in order to maximize

$$\ln(1 - \tau(s^1)) + \lambda \bar{w}\tau(s^1) + \ln(1 - \tau(s^2)) + \lambda \bar{w}\tau(s^2)$$

subject to the collective incentive compatibility constraint,

$$k - \theta_1 \geq \ln(1 - \tau(s^1)) - \ln(1 - \tau(s^2)) + \frac{\bar{w}}{w_2}(\tau(s^2) - \tau(s^1)).$$

The proof that  $\tau(s^1)$  is lower and that  $\tau(s^2)$  is higher, as compared to a model without collective incentive constraints, involves the following steps: First, use a Lagrangean approach, where  $\mu$  is the non-negative

multiplier on the collective incentive constraint. Second, derive first-order conditions and observe that  $\tau(s^1)$  decreases in  $\mu$  and that  $\tau(s^2)$  increases in  $\mu$ . Hence, if the collective incentive constraint is binding,  $\mu$  is strictly positive and  $\tau(s^1)$  is therefore larger and  $\tau(s^2)$  smaller than otherwise. Finally, it follows from the feasibility constraints that this also implies that  $\alpha(s^1)$  is smaller and that  $\alpha(s^2)$  is larger than otherwise if the collective incentive constraint is binding. ■

### Proof of Proposition 6

To prove the Proposition, we provide an example of a partial lie that satisfies the collective incentive compatibility constraints, for all  $\pi \in \Pi^c$ , but violates it for some  $\pi$  not in  $\Pi^c$ . Suppose that there are two possible taste parameters  $\Theta = \{\theta_1, \theta_2\}$  and two possible skill levels  $W = \{w_1, w_2\}$ .

We focus on a subset of states  $S' = \{s^1, s^2, s^3, s^4\}$ , which satisfy the following properties: In states  $s^1$  and  $s^2$ , all low-skilled individuals have a low valuation of the public good. In states  $s^3$  and  $s^4$ , all low-skilled individuals have a high valuation of the public good. In states  $s^2$  and  $s^3$ , all high-skilled individuals have a low valuation of the public good. In states  $s^1$  and  $s^4$ , they all have a high valuation of the public good.

Denote by  $x_{kl}(s)$  the payoff of an individual with taste parameter  $\theta_l$  and skill level  $w_k$  in state  $s$ . Suppose that the social choice function (and hence the truthful equilibrium of a direct mechanism implementing it) has the following properties: For some  $\epsilon > 0$ ,  $x_{11}(s^1) = x_{11}(s^4) + \epsilon$ ,  $x_{21}(s^1) = x_{21}(s^4) + \epsilon$ ,  $x_{12}(s^1) = x_{12}(s^4) - 2\epsilon$ , and  $x_{22}(s^1) = x_{22}(s^4) - 2\epsilon$ . In words, moving from the outcome for state  $s^1$  to the outcome for state  $s^4$  harms the low-skilled a little and makes the high-skilled substantially better off. Similar effects occur with reverse sign as one switches from the outcome for state  $s^3$  to the one for state  $s^4$ :  $x_{12}(s^3) = x_{12}(s^4) + \epsilon$ ,  $x_{22}(s^3) = x_{22}(s^4) + \epsilon$ ,  $x_{11}(s^3) = x_{11}(s^4) - 2\epsilon$ , and  $x_{21}(s^3) = x_{21}(s^4) - 2\epsilon$ .

Now consider an incomplete information type space which is such that all types assume that states  $s^1$  and  $s^3$  occur with probability  $\frac{1}{2}$  each, whereas states  $s^2$  and  $s^4$  occur with probability 0. Consider a partial lie by all high-skilled and low-skilled individuals which is such that all individuals, irrespective of their true taste parameter, announce a high taste parameter. The impact of this is as follows: In states  $s^1$  and  $s^3$  the outcome stipulated for state  $s^4$  is implemented. If the true state turns out to be  $s^1$ , low-skilled individuals with a low taste parameter are made worse off and high-skilled individuals with a high taste parameter are made better off by this deviation. This pattern is reversed if the true state turns out to be  $s^3$ . From an ex interim perspective, all individuals who are supposed to deviate benefit from this. Hence, collective incentive compatibility fails.

However, this deviation is not attractive under conditions of complete information. Suppose, for instance, that all individuals believe that  $s^1$  is the true state of the economy with probability 1, then low-skilled individuals with a low taste parameter do no longer benefit from this deviation. ■

### Proof of Proposition 7

The constraints in (18) trivially imply that the inequalities (20) and (19) hold. For instance, if (19) was violated and the true state of the economy was  $s = (s_{-k}, s_k^l)$ , individuals with taste parameter  $\theta_l$  would collectively lie about their taste parameters and induce an announced distribution  $(s_{-k}, s_k^{l+1})$ .

We now want to show that the inequalities (20) and (19) imply that the constraints in (18) are satisfied. First note that for given  $k$ ,  $s_{-k}$  and  $l$ , adding the inequalities (19) and (20), where the latter is formulated for  $l+1$  instead of  $l$ , yields

$$v(q(s_k^{l+1}, s_{-k}), \theta_{l+1}) - v(q(s_k^l, s_{-k}), \theta_{l+1}) \geq v(q(s_k^{l+1}, s_{-k}), \theta_l) - v(q(s_k^l, s_{-k}), \theta_l)$$

Using the assumption that  $v_q(q, \theta)$  is, for every given  $q$ , increasing in  $\theta$  makes it possible to show that

$$q(s_{-k}, s_k^{l+1}) \geq q(s_{-k}, s_k^l). \quad (34)$$

Second, suppose that the true state of the economy is  $(s_{-k}, s_k^l)$ . Then any manipulation of individuals with  $w^i = w_k$  requires the participation of those with  $\theta^i = \theta_l$ . If those individuals do not participate and reveal their characteristics truthfully, any alternative manipulation would induce a result that is incompatible with  $S_k$ , i.e., that occurs with probability zero. We may hence assume, without loss of generality, that the mechanism deters such manipulations.

Third, the inequalities (20) and (19) exclude that individuals with  $\theta^i = \theta_l$  participate in a “local deviation”, i.e., a manipulation that induces the outcome for state  $(s_{-k}, s_k^{l+1})$  or state  $(s_{-k}, s_k^{l-1})$  if the true state is  $(s_{-k}, s_k^l)$ . It remains to be shown that these individuals do also refuse participation for any other manipulation.

Consider a manipulation that attempts to induce the outcome for state  $(s_k^{l+j}, s_{-k})$  for some  $j > 1$ . We want to show that

$$v(q(s_k^l, s_{-k}), \theta_l) + V_k(s_k^l, s_{-k}) \geq v(q(s_k^{l+j}, s_{-k}), \theta_l) + V_k(s_k^{l+j}, s_{-k}). \quad (35)$$

The monotonicity of  $q$ , (see the inequality in (34)) and the assumption that  $v_q(q, \theta)$  is increasing in  $\theta$ , have the following implication: The local incentive constraint

$$v(q(s_k^{l+j-1}, s_{-k}), \theta_{l+j-1}) + V_k(s_k^{l+j-1}, s_{-k}) \geq v(q(s_k^{l+j}, s_{-k}), \theta_{l+j-1}) + V_k(s_k^{l+j}, s_{-k}).$$

implies that for all  $g \leq l + j - 1$ ,

$$v(q(s_k^{l+j-1}, s_{-k}), \theta_g) + V_k(s_k^{l+j-1}, s_{-k}) \geq v(q(s_k^{l+j}, s_{-k}), \theta_g) + V_k(s_k^{l+j}, s_{-k}).$$

A repeated application of this argument implies that

$$\begin{aligned} v(q(s_k^l, s_{-k}), \theta_l) + V_k(s_k^l, s_{-k}) &\geq v(q(s_k^{l+1}, s_{-k}), \theta_l) + V_k(s_k^{l+1}, s_{-k}) \\ &\geq v(q(s_k^{l+2}, s_{-k}), \theta_l) + V_k(s_k^{l+2}, s_{-k}) \\ &\geq \dots \end{aligned}$$

which establishes (35). A symmetric argument can be used to show that

$$v(q(s_k^l, s_{-k}), \theta_l) + V_k(s_k^l, s_{-k}) \geq v(q(s_k^{l-j}, s_{-k}), \theta_l) + V_k(s_k^{l-j}, s_{-k}),$$

for all  $j > 1$ . ■

## Proof of Proposition 8

The construction of a Groves mechanism is based on two observations. First, a surplus-maximizing choice of  $q$  implies that, for all  $k$ , for all  $s_{-k}$  and for all  $s_k^l \in S_k$ ,

$$\begin{aligned} s_k^l \in \operatorname{argmax}_{s_k^j \in S_k} & \sum_{g \neq k} f_g \sum_{h=1}^m s_{gh} v(q^*(s_k^j, s_{-k}), \theta_h) \\ & + f_k \sum_{h=1}^m s_{kh} v(q^*(s_k^j, s_{-k}), \theta_h) - r(q^*(s_k^j, s_{-k})). \end{aligned} \quad (36)$$

Second, incentive compatibility requires that for all  $k$ , for all  $s_{-k}$  and for all  $s_k^l \in S_k$ ,

$$s_k^l \in \operatorname{argmax}_{s_k^j \in S_k} v(q^*(s_k^j, s_{-k}), \theta_l) + V_k(s_k^j, s_{-k}). \quad (37)$$

A Groves mechanism is based on the idea to choose  $V_k$  such that two maximization problems are equivalent. This requires that (i) for every  $k$ , for all  $s_{-k}$ , for all  $s_k^l$ , and for all  $s_k^j$  the objective functions of the maximization problems (36) and (37) are equal,

$$\begin{aligned} V_k(s_k^j, s_{-k}) &= -v(q^*(s_k^j, s_{-k}), \theta_l) + \sum_{g \neq k} f_g \sum_{h=1}^m s_{gh} v(q^*(s_k^j, s_{-k}), \theta_h) \\ &\quad + f_k \sum_{h=1}^m s_{kh}^l v(q^*(s_k^j, s_{-k}), \theta_h) - r(q^*(s_k^j, s_{-k})) \end{aligned} \quad (38)$$

and (ii) that the right-hand side of (38) is indeed independent of  $s_k^l$ , the true distribution of preferences among individuals with skill level  $w_k$ . Generally, it is not possible to achieve (i) and (ii) simultaneously. However, if we impose the assumptions that skills are uniformly distributed ( $f_k = \frac{1}{n}$ , for each  $k$ ) and that there is no heterogeneity in public goods preferences among individuals with the same taste parameter (for each  $k$ ,  $s_k^1 = (f_k, 0, \dots, 0)$ ,  $s_k^2 = (0, f_k, 0, \dots, 0)$ ,  $\dots$ ,  $s_k^m = (0, \dots, 0, f_k)$ ) a Groves mechanism works. To see this, note first that (36) can now be written as follows: for every  $k$ , for every  $s_{-k}$  and for all  $s_k^l$ ,

$$s_k^l \in \operatorname{argmax}_{s_k^j \in S_k} \sum_{g \neq k} v(q^*(s_k^j, s_{-k}), \theta(s_g)) + v(q^*(s_k^j, s_{-k}), \theta(s_k^l)) - nr(q^*(s_k^j, s_{-k})), \quad (39)$$

where  $\theta(s_g)$  is the taste parameter of individuals with  $w^i = w_g$  in state  $s$ . If we now rewrite (38) accordingly, this yields

$$V_k(s_k^j, s_{-k}) = \sum_{g \neq k} v(q^*(s_k^j, s_{-k}), \theta(s_g)) - nr(q^*(s_k^j, s_{-k})) \quad (40)$$

which is just the standard definition of a Groves mechanism. ■

# CESifo Working Paper Series

for full list see [www.cesifo-group.org/wp](http://www.cesifo-group.org/wp)

(address: Poschingerstr. 5, 81679 Munich, Germany, [office@cesifo.de](mailto:office@cesifo.de))

---

- 2641 Martin Halla, Mario Lackner and Friedrich G. Schneider, An Empirical Analysis of the Dynamics of the Welfare State: The Case of Benefit Morale, May 2009
- 2642 Balázs Égert, Infrastructure Investment in Network Industries: The Role of Incentive Regulation and Regulatory Independence, May 2009
- 2643 Christian Gollier, Expected Net Present Value, Expected Net Future Value, and the Ramsey Rule, May 2009
- 2644 Sören Blomquist and Håkan Selin, Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates, May 2009
- 2645 Dominique Demougin, Oliver Fabel and Christian Thomann, Implicit vs. Explicit Incentives: Theory and a Case Study, May 2009
- 2646 Francesco C. Billari and Vincenzo Galasso, What Explains Fertility? Evidence from Italian Pension Reforms, May 2009
- 2647 Kjell Arne Brekke, Karen Evelyn Hauge, Jo Thori Lind and Karine Nyborg, Playing with the Good Guys – A Public Good Game with Endogenous Group Formation, May 2009
- 2648 Guglielmo Maria Caporale and Luis A. Gil-Alana, Multi-Factor Gegenbauer Processes and European Inflation Rates, May 2009
- 2649 Henning Bohn, A Static Model for Voting on Social Security, May 2009
- 2650 Markus Haavio and Kaisa Kotakorpi, The Political Economy of Sin Taxes, May 2009
- 2651 Augusto de la Torre, María Soledad Martínez Pería and Sergio L. Schmukler, Drivers and Obstacles to Banking SMEs: The Role of Competition and the Institutional Framework, May 2009
- 2652 Tobias Lindhe and Jan Södersten, Dividend Taxation, Share Repurchases and the Equity Trap, May 2009
- 2653 Assaf Razin and Edith Sand, Migration-Regime Liberalization and Social Security: Political-Economy Effect, May 2009
- 2654 Yin-Wong Cheung and Hiro Ito, A Cross-Country Empirical Analysis of International Reserves, May 2009
- 2655 Bart Cockx and Bruno Van der Linden, Flexicurity in Belgium. A Proposal Based on Economic Principles, May 2009

- 2656 Michael Melvin, Lukas Menkhoff and Maik Schmeling, Exchange Rate Management in Emerging Markets: Intervention via an Electronic Limit Order Book, May 2009
- 2657 Susanne Neckermann, Reto Cueni and Bruno S. Frey, What is an Award Worth? An Econometric Assessment of the Impact of Awards on Employee Performance, May 2009
- 2658 Steven Brakman, Harry Garretsen and Charles van Marrewijk, Economic Geography within and between European Nations: The Role of Market Potential and Density across Space and Time, May 2009
- 2659 Giovanni Facchini and Cecilia Testa, Reforming Legislatures: Is one House better than two?, May 2009
- 2660 Carsten Kowalczyk and Raymond Riezman, Trade Agreements, May 2009
- 2661 Oliver Falck, Stephan Heblich and Elke Luedemann, Identity and Entrepreneurship, May 2009
- 2662 Christian Lessmann and Gunther Markwardt, One Size Fits All? Decentralization, Corruption, and the Monitoring of Bureaucrats, May 2009
- 2663 Felix Bierbrauer, On the Legitimacy of Coercion for the Financing of Public Goods, May 2009
- 2664 Alessandro Cigno, Agency in Family Policy: A Survey, May 2009
- 2665 Claudia M. Buch and Christian Pierdzioch, Low Skill but High Volatility?, May 2009
- 2666 Hendrik Jürges, Kerstin Schneider, Martin Senkbeil and Claus H. Carstensen, Assessment Drives Learning: The Effect of Central Exit Exams on Curricular Knowledge and Mathematical Literacy, June 2009
- 2667 Eric A. Hanushek and Ludger Woessmann, Schooling, Cognitive Skills, and the Latin American Growth Puzzle, June 2009
- 2668 Ourania Karakosta, Christos Kotsogiannis and Miguel-Angel Lopez-Garcia, Does Indirect Tax Harmonization Deliver Pareto Improvements in the Presence of Global Public Goods?, June 2009
- 2669 Aleksandra Riedl and Silvia Rocha-Akis, Testing the Tax Competition Theory: How Elastic are National Tax Bases in OECD Countries?, June 2009
- 2670 Dominique Demougin and Carsten Helm, Incentive Contracts and Efficient Unemployment Benefits, June 2009
- 2671 Guglielmo Maria Caporale and Luis A. Gil-Alana, Long Memory in US Real Output per Capita, June 2009

- 2672 Jim Malley and Ulrich Woitek, Productivity Shocks and Aggregate Cycles in an Estimated Endogenous Growth Model, June 2009
- 2673 Vivek Ghosal, Business Strategy and Firm Reorganization under Changing Market Conditions, June 2009
- 2674 Francesco Menoncin and Paolo M. Panteghini, Retrospective Capital Gains Taxation in the Real World, June 2009
- 2675 Thomas Hemmelgarn and Gaëtan Nicodème, Tax Co-ordination in Europe: Assessing the First Years of the EU-Savings Taxation Directive, June 2009
- 2676 Oliver Himmler, The Effects of School Competition on Academic Achievement and Grading Standards, June 2009
- 2677 Rolf Golombek and Michael Hoel, International Cooperation on Climate-Friendly Technologies, June 2009
- 2678 Martin Cave and Matthew Corkery, Regulation and Barriers to Trade in Telecommunications Services in the European Union, June 2009
- 2679 Costas Arkolakis, A Unified Theory of Firm Selection and Growth, June 2009
- 2680 Michelle R. Garfinkel, Stergios Skaperdas and Constantinos Syropoulos, International Trade and Transnational Insecurity: How Comparative Advantage and Power are Jointly Determined, June 2009
- 2681 Marcelo Resende, Capital Structure and Regulation in U.S. Local Telephony: An Exploratory Econometric Study; June 2009
- 2682 Marc Gronwald and Janina Ketterer, Evaluating Emission Trading as a Policy Tool – Evidence from Conditional Jump Models, June 2009
- 2683 Stephan O. Hornig, Horst Rottmann and Rüdiger Wapler, Information Asymmetry, Education Signals and the Case of Ethnic and Native Germans, June 2009
- 2684 Benoit Dostie and Rajshri Jayaraman, The Effect of Adversity on Process Innovations and Managerial Incentives, June 2009
- 2685 Peter Egger, Christian Keuschnigg and Hannes Winner, Incorporation and Taxation: Theory and Firm-level Evidence, June 2009
- 2686 Chrysovalantou Milliou and Emmanuel Petrakis, Timing of Technology Adoption and Product Market Competition, June 2009
- 2687 Hans Degryse, Frank de Jong and Jérémie Lefebvre, An Empirical Analysis of Legal Insider Trading in the Netherlands, June 2009
- 2688 Subhasish M. Chowdhury, Dan Kovenock and Roman M. Sheremeta, An Experimental Investigation of Colonel Blotto Games, June 2009

- 2689 Alexander Chudik, M. Hashem Pesaran and Elisa Tosetti, Weak and Strong Cross Section Dependence and Estimation of Large Panels, June 2009
- 2690 Mohamed El Hedi Arouri and Christophe Rault, On the Influence of Oil Prices on Stock Markets: Evidence from Panel Analysis in GCC Countries, June 2009
- 2691 Lars P. Feld and Christoph A. Schaltegger, Political Stability and Fiscal Policy – Time Series Evidence for the Swiss Federal Level since 1849, June 2009
- 2692 Michael Funke and Marc Gronwald, A Convex Hull Approach to Counterfactual Analysis of Trade Openness and Growth, June 2009
- 2693 Patricia Funk and Christina Gathmann, Does Direct Democracy Reduce the Size of Government? New Evidence from Historical Data, 1890-2000, June 2009
- 2694 Kirsten Wandschneider and Nikolaus Wolf, Shooting on a Moving Target: Explaining European Bank Rates during the Interwar Period, June 2009
- 2695 J. Atsu Amegashie, Third-Party Intervention in Conflicts and the Indirect Samaritan's Dilemma, June 2009
- 2696 Enrico Spolaore and Romain Wacziarg, War and Relatedness, June 2009
- 2697 Steven Brakman, Charles van Marrewijk and Arjen van Witteloostuijn, Market Liberalization in the European Natural Gas Market – the Importance of Capacity Constraints and Efficiency Differences, July 2009
- 2698 Huifang Tian, John Whalley and Yuezhou Cai, Trade Sanctions, Financial Transfers and BRIC's Participation in Global Climate Change Negotiations, July 2009
- 2699 Axel Dreher and Justina A. V. Fischer, Government Decentralization as a Disincentive for Transnational Terror? An Empirical Analysis, July 2009
- 2700 Balázs Égert, Tomasz Koźluk and Douglas Sutherland, Infrastructure and Growth: Empirical Evidence, July 2009
- 2701 Felix Bierbrauer, Optimal Income Taxation and Public Goods Provision in a Large Economy with Aggregate Uncertainty, July 2009