

Bierbrauer, Felix

Working Paper

Optimal income taxation, public goods provision and robust mechanism design

Preprints of the Max Planck Institute for Research on Collective Goods, No. 2008,31

Provided in Cooperation with:

Max Planck Institute for Research on Collective Goods

Suggested Citation: Bierbrauer, Felix (2008) : Optimal income taxation, public goods provision and robust mechanism design, Preprints of the Max Planck Institute for Research on Collective Goods, No. 2008,31, Max Planck Institute for Research on Collective Goods, Bonn

This Version is available at:

<https://hdl.handle.net/10419/32223>

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.

**Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2008/31**



**Optimal Income Taxation,
Public Goods Provision
and Robust Mechanism
Design**

Felix Bierbrauer



MAX PLANCK SOCIETY



Optimal Income Taxation, Public Goods Provision and Robust Mechanism Design

Felix Bierbrauer

August 2008

Optimal Income Taxation, Public Goods Provision and Robust Mechanism Design

Felix Bierbrauer*

MPI, Bonn

August 4, 2008

Abstract

This paper extends the model of optimal income taxation due to Mirrlees (1971) and includes private information on public goods preferences. A mechanism design approach is used to establish the following result: If policies are required to be robustly implementable in the sense of Bergemann and Morris (2005), then the optimality conditions in the extended model with uncertainty about tax and expenditure policies, are the same as in the standard model of optimal income taxation. The paper provides a foundation for a widely used assumption in public finance, namely that individuals optimize their behaviour subject to a predetermined and commonly known tax system.

Keywords: Optimal Taxation, public goods provision, Revelation of Preferences.

JEL: D71, D82, H21, H41

1 Introduction

This paper studies income taxation and public goods provision under the assumption that individuals have private information about their skills and about their preferences for a public good. Traditionally, normative public economics has separated the analysis of tax systems and the revelation of preferences for public goods. Most of the literature on optimal taxation ignores public goods and characterizes an optimal tax system under the assumption that the government has an exogenously given funding requirement. Papers on income taxation which include public goods provision assume that

*I am very grateful for numerous discussions with Martin Hellwig. I also benefited from conversations with Stefanie Brilon, Christoph Engel, Mike Golosov, and Marco Sahm. I thank seminar participants at the Massachusetts Institute of Technology, the Max Planck Institute for Research on Collective Goods, Bonn, and participants at the Public Economic Theory Meeting in Nashville 2007 and the European Meeting of the Econometric Society in Budapest 2007. I also thank the Massachusetts Institute of Technology for its hospitality.

preferences for public goods are common knowledge. They focus on how the distortions associated with taxation affect the level of public goods provision, under the premise that a question such as “how many individuals have a high, medium or low valuation of the public good” has already somehow been answered.¹

By contrast, the literature on public goods provision under asymmetric information focusses on the problem how to determine the social benefit of a public good, given that preferences of individuals are not observable. Moreover, this literature does not allow for taxation as a source of funds.²

A joint treatment of these two problems is difficult because they are typically analyzed in very different models. The theory of optimal taxation assumes a large economy in which individuals act as “price-takers”, i.e., they take the tax system and the provision of public goods as given. The literature on the revelation of public goods preferences, by contrast, is based on a “small” economy in which each single individual has a noticeable impact on public goods provision.

This paper extends the large economy model of the theory of optimal income taxation so as to allow for incomplete information about the cross-section distribution of preferences for a public good. Hence, it will be necessary to acquire information from individuals to determine the social benefits from public goods provision. In particular, this implies that individuals face uncertainty about the tax and expenditure policy. The main question is what differences, if any, the privateness of information on public goods preferences makes for optimal taxation and public goods provision.

The analysis uses a mechanism design approach. Each individual is asked to send a message. The mechanism determines the public good provision level and the tax system as a function of the profile of messages that it has received. By the Revelation Principle an individual’s message may be taken to consist of the individual’s public goods preference and the individual’s skill level.

The set of implementable policies depends on the probabilistic beliefs of individuals about the public good provision level and the shape of the tax system. Consequently, alternative assumptions about individual beliefs lead to alternative tax and expenditure policies. Given that there is no way of telling what “reasonable” beliefs are, this notion of implementability is problematic. I therefore invoke the robustness requirement of Bergemann and Morris (2005) so that a social choice function has to be implementable for *any* set of beliefs that individuals might have.

Given this robustness requirement, I extend the *Taxation Principle*, due to Hammond (1979) and Guesnerie (1995), to show that any mechanism is equivalent to a system specifying a public good provision level and an income tax schedule as a function of the cross-section distribution of preferences, which is such that, given the public good

¹Examples are Atkinson and Stern (1974), Wilson (1991), Boadway and Keen (1993), Nava et al. (1996), Sandmo (1998), Gaube (2000), Hellwig (2004), Kaplow (2006) or Gahvari (2006).

²Examples of this approach include Clarke (1971), Groves (1973), d’Aspremont and Gérard-Varet (1979), Mailath and Postlewaite (1990), Hellwig (2003), or Neeman (2004).

provision level and the income tax schedule, each individual chooses labor supply to maximize utility. No further incentive compatibility condition, e.g., on the communication of public goods preferences is involved.

The *extended Taxation Principle* has the following implication: the model with uncertainty about optimal policies gives rise to the same optimality conditions for income taxation and public goods provision as the existing literature which is based on a deterministic environment. The main finding can therefore be summarized as follows: Given the robustness requirement of Bergemann and Morris (2005), the optimal tax and expenditure policy for an economy with incomplete information on the desirability of public goods provision, may be derived under the hypothesis that there is complete information and that individuals behave optimally subject to a predetermined income tax system and a predetermined public goods provision level.

The argument proceeds as follows: Attention is limited to social choice functions that can be implemented as outcomes of anonymous games in a large economies. The requirement of anonymity implies that all individuals with the same characteristics are treated equally by the allocation mechanism. By an adaption of a result of Bergemann and Morris (2005) I find that a social choice function is robustly implementable if and only if it is incentive compatible from an ex post perspective, i.e., the mechanism which implements the social choice function must be such that, no individual has an incentive to revise the own action after the actions of other individuals have been revealed.

From an ex post perspective, the distribution of preferences is known. Consequently, robust mechanism design gives rise to the same implementability conditions as a model with complete information on the distribution of preferences. For a complete information environment, the *Taxation Principle* says that that any incentive compatible allocation can be decentralized via a suitably chosen income tax system. Instead of relying on a game where individuals send messages one may use a system where individuals choose their labor supplies under a given tax system.

These observations imply the *extended Taxation Principle* that is derived in this paper: an admissible social choice function can be *represented* by a rule which specifies a public good provision level and an income tax schedule as a function of the distribution of public goods preferences. Whatever the distribution of public goods preferences, the tax schedule has to be such that individual labor supply is utility maximizing. To implement this outcome it is no longer possible to use a mechanism where a choice of labor supply is the only action that is taken by individuals because the information on public goods preferences needs to be transmitted to the policy maker. However, this has no bearing on incentive compatible policies. In a large economy, no individual has a noticeable impact on public good provision so that individuals are trivially willing to reveal their public goods preferences.

The remainder of the paper is organized as follows. The Model is introduced in Section 2. Section 3 contains the characterization of robustly implementable tax and expen-

diture policies. In Section 4, I briefly discuss some recent related literature. The last Section contains concluding Remarks. All proofs are in the Appendix.

2 The Model

A workhorse of public economics is the model of optimal income taxation due to Mirrlees (1971). In the following, I introduce an extended version of this model with uncertainty about the social benefits of public goods provision.

Suppose that there is a continuum of individuals identified with the unit interval $I = [0, 1]$. The utility function of individual i is given by

$$u^i = u(q, c, y, w^i, \theta^i),$$

where q is the amount of a public good, c is the individual's consumption of private goods, or, after-tax income, and y is the individual's contribution to the economy's output, or, pre-tax income. It is assumed that the utility function u increases in q and c and decreases in y .

w^i is a skill parameter that affects how much effort an individual has to exert in order to produce an additional unit of output. For each i , w^i belongs to a finite ordered set $W = \{w_1, \dots, w_n\}$. θ^i is a taste parameter that affects an individual's valuation of the public good and belongs to a finite ordered set $\Theta = \{\theta_1, \dots, \theta_m\}$.

A standard assumption in the theory of optimal income taxation is that the single crossing property holds, i.e., at any point in a $y - c$ diagram, the slope of an individual's indifference curve is a decreasing function of the skill parameter,

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

for any k , any q , and any $\theta \in \Theta$. u_c and u_y are the partial derivatives of u with respect to c and y , respectively. The interpretation is that a high-skilled individual needs to exert less effort in order to produce additional output and hence needs less compensation for a marginal increase in the output requirement.

A standard assumption in the theory of public goods under incomplete information is that an individual's willingness to give up private goods for an increase in public good provision is an increasing function of the taste parameter. Consequently, for any given q and c ,

$$-\frac{u_q(q, c, y, w, \theta_l)}{u_c(q, c, y, w, \theta_l)} < -\frac{u_q(q, c, y, w, \theta_{l+1})}{u_c(q, c, y, w, \theta_{l+1})}, \quad (2)$$

for any l , any y , and any $w \in W$.

For instance, if preferences take the additively separable form

$$u(q, c, y, w, \theta) = v(q, \theta) + V(c, y, w), \quad (3)$$

then these properties hold if the function V satisfies the single crossing property and the function v is such that the marginal utility that an individual derives from the

public good is increasing in θ , $v_q(q, \theta_l) < v_q(q, \theta_{l+1})$.

The informational structure is as follows. Individuals have private information about their skill levels and their taste parameters. There is a commonly known cross-section distribution of skill levels $f = (f_1, \dots, f_n)$, where f_k is the population share of individuals with skill level w_k . Hence, regarding the distribution of skill parameters, there is no uncertainty at the aggregate level.

Uncertainty about the distribution of public goods preferences arises via the joint distribution of taste and productivity parameters. This distribution depends on the state of the economy which is taken to be the realization of a random variable \tilde{s} that takes values in a set S . For a given state $s \in S$, let $s_k = (s_{k1}, \dots, s_{km})$ be the distribution of taste parameters among individuals with skill level w_k , i.e., s_{kl} is the fraction of individuals with skill level w_k who have a taste parameter θ_l . A state of the economy s is characterized by such a conditional distribution for each skill level and can be written as $s = (s_k)_{k=1}^n$.

To give an example, let $W = \{w_1, w_2\}$ and $\Theta = \{\theta_1, \theta_2\}$. The information structure is such that the fractions f_1 and f_2 of low-skilled and high-skilled individuals in the population are common knowledge. Uncertainty about the state of the economy arises because the fraction of high-skilled individuals with a high valuation of the public good, s_{22} , and the fraction of low-skilled individuals with a high valuation of the public good, s_{12} , are random quantities.

Existing models in the theory of optimal income taxation can be viewed as special cases of this environment which have in common that there is no uncertainty about public goods preferences. For instance, Atkinson and Stern (1974) consider a model with a representative individual whose utility function is known. Boadway and Keen (1993) or Gahvari (2006) consider a model in which individuals differ only in their productive abilities. If an individual's skill level is known, then the utility function of the individual and, in particular, his preference for the public good, is known as well. Hence, public goods preferences are not a genuine source of heterogeneity among individuals. Moreover, at an aggregate level, since the distribution f of skill parameters in the population is known, the distribution of public goods preferences is also known. This implies that there is no problem of information aggregation that complicates the decision on an optimal supply of public goods.

I use a mechanism design approach to characterize the allocations, or social choice functions that can be implemented. A social choice function consists of a provision rule for the public good $q : s \mapsto q(s)$ that specifies for each state s how much of the public good is provided, and a pair of functions $c : (s, w, \theta) \mapsto c(s, w, \theta)$ and $y : (s, w, \theta) \mapsto y(s, w, \theta)$, which determine for each state, each skill level, and each taste parameter a consumption level and an output requirement, respectively. A social choice function is said to

be feasible, if for every s ,

$$\sum_{k=1}^n f_k \sum_{l=1}^m (y(s, w_k, \theta_l) - c(s, w_k, \theta_l)) s_{kl} \geq r(q(s)) , \quad (4)$$

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

I limit attention to social choice functions that can be implemented by an anonymous game.³ In the present setting, anonymity means that the decision on public good provision is left unaffected by a permutation of individual actions. For consumption levels and output requirements, anonymity holds provided that all individuals who choose the same action get the same $y - c$ bundle, and that the menu of available bundles is also unaffected by a permutation of actions. The interpretation is that attention is limited to institutional arrangements which satisfy an *equal treatment* condition.

Formally, an anonymous allocation mechanism $M = (A, Q, C, Y)$ specifies an action set A and a collection of functions, where $Q : \alpha \mapsto Q(\alpha)$ gives the level of public good provision as a function of the cross-section distribution of actions α , and the functions $C : (\alpha, a) \mapsto C(\alpha, a)$ and $Y : (\alpha, a) \mapsto Y(\alpha, a)$ specify a consumption level and an output requirement as a function of an individual's message a and of the distribution of messages α .

A social choice function is called *implementable* if there is some anonymous mechanism such that, in every state s , the equilibrium allocation is equal to the allocation stipulated by the social choice function. Formally, suppose that $\sigma^* : W \times \Theta \rightarrow A$ is an equilibrium strategy for the game induced by M and let $\alpha^*(s)$ be the cross-section distribution of actions induced by strategy σ^* in state s . The mechanism M is said to implement the social choice function (q, c, y) , if for all s ,

$$Q(\alpha^*(s)) = q(s)$$

and

$$C(\alpha^*(s), \sigma^*(w, \theta)) = c(s, w, \theta) \quad \text{and} \quad Y(\alpha^*(s), \sigma^*(w, \theta)) = y(s, w, \theta) ,$$

for all (w, θ) .

The set of implementable social choice functions depends on the equilibrium concept that is used. In the following section I will first provide a characterization of the social choice functions that are robustly implementable in the sense of Bergemann and Morris (2005). I will then prove a *Taxation Principle* which relates the set of implementable social choice functions to the set of outcomes that can be reached with an income tax system.

³For a more extensive discussion of anonymous games, see Kalai (2004).

3 Robust Mechanism Design and the Taxation Principle

3.1 Robust Mechanism Design

For the given environment, a “natural” solution concept is as follows: a social choice function is said to be implementable if it can be made the equilibrium allocation of a game of incomplete information, i.e., individual behaviour must be optimal from an *interim* perspective where individuals know their own characteristics but have only probabilistic beliefs about the characteristics and the behavior of others. Following Bergemann and Morris (2005), I refer to this as implementability as *an interim Nash equilibrium*.⁴

A difficulty with the notion of implementability as an interim Nash equilibrium is that it requires assumptions on individual beliefs about the state \tilde{s} of the economy and possibly also about higher order beliefs, i.e., beliefs about the beliefs of other individuals, etc. Given that any specific assumption on the beliefs of individuals is hard to defend, I follow Bergemann and Morris (2005) and require that the set of possible outcomes should be robust with respect to the specification of individual beliefs, i.e., a social choice function is required to be implementable as an interim Nash equilibrium for any set of beliefs that individuals may have.

An alternative solution concept is implementability as an ex post equilibrium. This requires that there is an equilibrium σ^* such that after the distribution of actions by all other players has been revealed, no player would be willing to choose an action that differs from the one initially chosen.⁵ Formally, σ^* is an ex post equilibrium for the game induced by mechanism $M = (A, Q, C, Y)$ if, for all $s \in S$, and for all $(w, \theta) \in W \times \Theta$,

$$\sigma^*(w, \theta) \in \operatorname{argmax}_{a \in A} u(Q(\alpha^*(s)), C(\alpha^*(s), a), Y(\alpha^*(s), a), w, \theta).$$

For a given set of individual beliefs, this notion is stronger than the requirement of implementability as an interim Nash equilibrium. It requires that each individual has an action that is a best response for each possible state of the economy. An interim Nash equilibrium requires only that the equilibrium action is a best response “on average”, when forming expectations about \tilde{s} . Hence, the notion of an ex post equilibrium may be criticized because it gives a too restrictive view on the tax and expenditure policies that can be achieved under conditions of private information on skill levels and public goods preferences.⁶

However, the main result of Bergemann and Morris (2005) is that the set of robustly

⁴If one adds the assumption that individual beliefs are derived according to Bayes’ rule from some common prior distribution this yields the special case of implementability as a *Bayes-Nash equilibrium*.

⁵In the given environment, implementability in an ex post equilibrium is equivalent to implementability in a dominant strategy equilibrium. This equivalence holds in models with *private values*, where the private information of some individual i is not directly payoff relevant for some other individual j . See Bergemann and Morris (2005) for details.

⁶For instance, the Gibbard-Satterthwaite theorem provides conditions such that only dictatorial social choice functions are implementable as an ex post equilibrium. This however requires that there is no domain restriction on preferences; see Samuelson (1967) for a discussion of this requirement.

implementable social choice functions and the set of ex post implementable social choice functions coincide, i.e., if we allow for mechanisms that exploit the belief structure of individuals, but at the same time insist that implementation is possible whatever those beliefs are, then we can not enlarge the set of implementable social choice functions relative to those that are implementable as an ex post equilibrium.

A formal derivation of this result for the large economy model in this paper is relegated to the Appendix because it requires a substantial investment in notation.⁷ The main logic is as follows. If a social choice function is implementable for any set of beliefs, then, in particular, it must be implementable in a “complete information environment” where all individuals “know” (assign probability 1 to) the true state of the economy. This implies that a robustly implementable social choice function is also implementable as an ex post equilibrium. The converse implication holds because implementability as an ex post equilibrium means that a social choice function can be made the outcome of a game in which individuals have an action that is a best response in each possible state s . Trivially, this implies that this action is also a best response if the same game form is used to implement the social choice function, but individuals have incomplete information and take expectations about \tilde{s} .

3.2 Incentive Compatible Social Choice Functions

Given the equivalence between robustly implementable social choice functions and those that are implementable as an ex post equilibrium, I will in the following focus on the latter. Proposition 1 and Corollary 1 below clarify how the requirement of implementability as an ex post equilibrium restricts the set of social choice functions.

Proposition 1 Suppose that a mechanism $M = (A, Q, C, Y)$ implements the social choice function (q, c, y) in an ex post equilibrium. Then, a mechanism $M' = (A', Q', C', Y')$ with $A' = W \times \Theta$ which is such that, for all s ,

$$Q'(s) = q(s)$$

and

$$C'(s, w, \theta) = c(s, w, \theta) \quad \text{and} \quad Y'(s, w, \theta) = y(s, w, \theta),$$

for all (w, θ) , has an ex post equilibrium σ' , where $\sigma'(\theta, w) = (\theta, w)$, for all (θ, w) .

Proposition 1 is a Revelation Principle for a large economies: if there exists some anonymous game with the property that the equilibrium allocation coincides with the social choice function, then there is also a direct mechanism, i.e., a mechanism such that the action set A is equal to the set of types $W \times \Theta$, that reaches the social choice

⁷In particular, one has to enrich the definition of a type in such a way that this includes not only the individual’s skill level and taste parameter but also the individual’s beliefs about the types of other individuals.

function in an equilibrium where all individuals follow a strategy of truth-telling.⁸ A more concise way of stating this observation is given in the following Corollary.

Corollary 1 A social choice function (q, c, y) is implementable in an ex post equilibrium if and only if it satisfies the following *incentive compatibility constraints*: For all s , for all (θ, w) , and for all $(\hat{\theta}, \hat{w})$,

$$u(q(s), c(s, w, \theta), y(s, w, \theta), w, \theta) \geq u(q(s), c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \hat{\theta}), w, \theta). \quad (5)$$

The incentive compatibility constraints are interpreted as follows: Individuals choose their announcements to the mechanism designer ex interim, i.e., after having learned their own characteristics but without knowing s . Now consider an ex post perspective, where the state s of the economy has been revealed to individuals, and suppose that after s has become known, individuals were given the opportunity to revise their announcement to the mechanism. If the inequalities in (5) hold, then, whatever the value of s , no individual would benefit from choosing an announcement that differs from truth-telling.

The large economy specification implies that an individual has no impact on the distribution of actions in the revelation game and hence on public good provision: $q(s)$ appears both on the left hand side and on the right hand side of (5). Individuals thus act as “price-takers,” in the sense of taking the state of the economy s that is perceived by the mechanism designer and hence the amount of the public good as given. This is a crucial difference to the literature on public goods in “small” economies where each individual can influence how much of a public good is provided.

The following Proposition establishes that if preferences are additively separable then incentive compatibility requires, in addition, that individuals who have the same skill level but differ in their public goods preferences are treated equally in terms of the consumption that they receive and the output they have to provide.

Proposition 2 Let $u(q, c, y, w, \theta) = v(q, \theta) + V(c, y, w)$. The social choice function (q, c, y) is ex post individually incentive compatible if and only if the following two properties are satisfied:

- i) Individuals who differ only in their public goods preferences but have identical skills are treated equally; i.e., for all s , for all w and for every pair θ and θ' ,

$$V(c(s, w, \theta), y(s, w, \theta), w) = V(c(s, w, \theta'), y(s, w, \theta'), w). \quad (6)$$

⁸There is a degree of freedom for the construction of the mechanism M' . It is required only that individuals prefer truth-telling in response to truth-telling of all other individuals. Hence, it can be left open what outcomes M' specifies in situations where the distribution of actions α' is such that there is no state $s \in S$ with $\alpha' = s$.

- ii) For any given taste parameter, an individual with skill level w , prefers the bundle $(c(s, \theta, w), y(s, \theta, w))$ over the bundle $(c(s, \theta, w'), y(s, \theta, w'))$; i.e., for all s , for all θ and for every pair w and w' ,

$$V(c(s, w, \theta), y(s, w, \theta), w) \geq V(c(s, w', \theta), y(s, w', \theta), w). \quad (7)$$

In the theory of optimal income taxation, the inequalities in (7) are known as self-selection constraints.⁹ They ensure that, for given θ , each skill type chooses the “right” $c - y$ pair from the menu $\{(c(s, w, \theta), y(s, w, \theta))\}_{w \in W}$

The equalities in (6) imply that taste parameters can not be used as a screening device. Individuals who differ only with respect to their preferences for the public good get the same private goods consumption and the same output requirement.¹⁰ Hence, the tax system cannot discriminate individuals according to their willingness to pay for public goods. The reason is that each individual takes the decision on public good provision as given and that an individuals’ marginal rate of substitution between consumption and leisure is not affected by the public good.

The crucial idea in the mechanism design literature on public goods provision in the tradition of Clarke (1971), Groves (1973), or d’Aspremont and Gérard-Varet (1979) is that payments are calibrated in such a way that individuals reveal their preferences. Proposition 2 and Corollary 1 show that this idea becomes inapplicable. In a large economy, individuals can not affect how much of the public good is provided. Moreover, with separable preferences, differences in public goods preferences can not be used as a screening device. Taken together, these two facts imply that the tax system can only exploit differences in productive abilities.

3.3 The extended Taxation Principle

I now turn to the main result of the paper. The following Proposition, which I refer to as the *extended Taxation Principle*, relates the social choice functions that can be reached via some arbitrary allocation mechanism to those that can be decentralized by means of a suitably chosen income tax system.

Proposition 3 A social choice function (q, c, y) satisfies the resource constraint (4) and is incentive compatible if and only if there exists a function $\tau : (s, y) \mapsto \tau(s, y)$ satisfying the following properties, for every $s \in S$:

- i) Consumption equals after tax income: for every (w, θ) ,

$$c(s, w, \theta) = y(s, w, \theta) - \tau(s, y(s, w, \theta)).$$

⁹See, for instance, Stiglitz (1982) or Boadway and Keen (1993).

¹⁰The proposition only says that utility levels are the same. For an efficient allocation, the joint utility level of individuals with θ and θ' is generated with minimal resources, so that also $(c(s, w, \theta), y(s, w, \theta)) = (c(s, w, \theta'), y(s, w, \theta'))$.

- ii) Individuals maximize utility subject to the given income tax schedule, i.e. for every (w, θ) ,

$$y(s, w, \theta) \in \operatorname{argmax}_y u(q(s), y - \tau(s, y), y, w, \theta). \quad (8)$$

- iii) Tax revenues are sufficient to cover the cost of public good provision,

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

Proposition 3 gives the implications of implementability as an ex post equilibrium for the analysis of tax systems. The postulate that individual behavior in a revelation game is a best response, for each possible state s , translates into the requirement that an individual's choice of pre-tax income y is utility maximizing given a tax system $\tau(s, \cdot)$ that depends on the state of economy. Put differently, individuals are assumed to choose y as if they knew the state s and therefore the income tax system that relates after-tax income to pre-tax income in state s .

In particular, this implies that the uncertainty about the state \tilde{s} and hence about public expenditures and the shape of the tax system do not affect individual behavior. For instance, risk attitudes do not affect the set of admissible tax systems because individuals are assumed to behave as if they knew the realization of \tilde{s} .

As a consequence, existing models of optimal income taxation with a commonly known state of the economy and the present model have the same constraints on the set of admissible policies, and lead to the same characterization of optimal tax systems and provision rules for the public good. The large economy specification implies that no individual can affect the state of the economy as perceived by the policy maker, i.e., individuals take s as given when undertaking their choices. As a consequence, there is no constraint such that the tax system and the expenditure policy chosen for some state s would restrict the set of admissible tax systems and expenditure policies for some other state s' . Hence, even if the distribution of public goods preferences \tilde{s} is a random quantity, an optimal social choice function requires that, for each possible realization $s \in S$, welfare is maximized subject to the feasibility constraint and the constraint that individual labor supply is utility maximizing given the tax system.

However, the models differ in the following sense: In the standard model without aggregate uncertainty, the mechanism that implements the social choice function can be such that individuals are confronted with a tax system and the choice of labor supply is the only action that they take. In the extended model by contrast, prior to the operation of the tax system the state of the economy has to be determined. This requires an initial phase of information aggregation that enables the mechanism designer to determine s . However, this need to learn s does not generate an incentive problem. Imposing that individual behavior is utility maximizing under the finally chosen tax system is both necessary and sufficient for incentive compatibility.

4 Discussion

The previous analysis has been based on a large economy model in which no single individual has a noticeable impact on the policy maker's perception of the state of his economy and hence on the policy that is implemented. Hence, there is no reason for individuals to communicate their public goods preferences in a strategic manner and incentive compatibility conditions refer only to the labor supply decisions of individuals.

However, individuals may also be willing to lie about their public goods preferences. Recall from Proposition 2 that, with additively separable preferences, individuals who have the same skills and differ only in their public goods preference are treated equally. This implies that, with a direct mechanism, individuals are willing to announce any taste parameter. Consequently, if s is the true distribution of public goods preferences, then every profile of taste announcements \hat{s} is consistent with individual actions being a best response from an ex post perspective.¹¹

This raises the question whether incentive compatibility is really a sufficient condition for the implementability of a social choice function. There are some recent papers in public economics that use a mechanism design approach and address this question. Below, I briefly discuss the different approaches and argue that they do not upset the main result of this paper, i.e., the standard deterministic model of optimal income taxation is robust in the sense that incomplete information per se does not imply that the characterization of optimal tax and expenditure policies needs to be revised.

Bassetto and Phelan (2008) and Kocherlakota and Phelan (2007) study problems of optimal taxation in a large economy with aggregate uncertainty. They are concerned with the problem that a direct revelation game may give rise to multiple equilibria.¹² They take an agnostic approach and assume that the policy maker has no device for making an equilibrium selection. As a consequence they propose to evaluate social choice functions using a maximin objective function so that the worst possible outcome should be made as attractive as possible.

Bierbrauer and Hellwig (2008) study public goods provision in a large economy with aggregate uncertainty and Bierbrauer (2007) studies optimal income taxation and public goods provision in such a model. These papers argue that imposing only incentive compatibility may give rise to implausible results. For instance, individuals may better off if a public good is not provided, but are still assumed to communicate their preferences in such a way that (infinitesimally) they increase the probability of provision. However, since it is also a best response to communicate preferences such that public good provision becomes less likely the assumption of truth-telling is problematic. To

¹¹We may assume that individuals will reveal their skills truthfully in any state s because of the self-selection constraints. Hence, multiplicity is an issue only with respect to the announced profile of public goods preferences.

¹²The general question whether it is possible to implement a social choice function as a unique equilibrium of some, possibly non-direct, mechanism has been studied before in the mechanism design literature; see Jackson (2001) or Moore (1992) for an overview.

avoid this kind of situation, Bierbrauer and Hellwig (2008) and Bierbrauer (2007) allow subcoalitions of individuals to coordinate on an equilibrium that is beneficial for them but has detrimental consequences from a welfare perspective, and develop a notion of implementation as a coalition-proof equilibrium.

However, mechanisms based on sampling provide a straightforward solution to these problems of equilibrium multiplicity on the one hand and multiple best responses on the other hand. Such a mechanism can implement a social choice function in a unique *ex post* equilibrium in which individuals have a unique best response.¹³ The idea is first to draw a random sample of individuals for the purpose of preference elicitation. Based on this information, the decision on tax and expenditure policies is taken, and the individuals who have not been in the sample choose their labor supply subject to the given tax system.

This brief discussion can be summarized as follows: the papers mentioned above emphasize that individuals can coordinate their reactions to public policy in a way that restricts the set of admissible social choice functions. Empirically, this is a very plausible restriction and it is certainly interesting to study the implications of this approach. However, the literature based on sampling mechanisms shows that equilibrium multiplicity is not an inevitable implication of an environment with aggregate uncertainty. This implies that one can determine the optimal tax and expenditure policy as in a standard public economics model with no aggregate uncertainty.

5 Concluding Remarks

The theory of optimal income taxation studies welfare maximizing tax policies for an environment where the distribution of abilities is known but individual abilities are unobservable. The theory of public goods under asymmetric information studies optimal allocations for an environment where the distribution of preferences is unknown.¹⁴ This paper has analyzed a large economy model that combines a known distribution of abilities with an unknown distribution of public goods preferences.

It has been found that an unknown distribution of public goods preferences is not a binding constraint. In a large economy, no individual can affect spending decisions that are based on an aggregate such as the cross-section distribution of preferences. Consequently, individuals have no incentive to hide their true public goods preference. In a system with millions of individuals the idea that any one individual may be pivotal for expenditure decisions and that this should be reflected in the individual's tax payment becomes moot.

However, uncertainty about preferences induces randomness into optimal tax and ex-

¹³The idea of a sampling mechanism has been introduced by Green and Laffont (1979). A more recent application is Gary-Bobo and Jaaidane (2000). Details on how to apply this idea to a model with both skill and taste heterogeneity can be found in Bierbrauer and Sahm (2008).

¹⁴The typical model has utility functions that are linear in money so that utility is perfectly transferable. As a consequence, knowing the cross-section distribution is sufficient for an optimal decision on public good provision, even if individual preferences are unknown.

penditure policies. We introduced the requirement that a policy mechanism must not be based on assumptions about the individuals' perception of this randomness. This notion of robustness was then shown to imply that policy mechanisms have to be designed as if individuals knew the true distribution of preferences. The analysis hence provides a foundation for a widely used assumption in public finance, namely that individuals optimize their behaviour subject to a given tax system.

References

- Atkinson, A. and Stern, N. (1974). Pigou, taxation and public goods. *Review of Economic Studies*, 41:119–128.
- Bassetto, M. and Phelan, C. (2008). Tax riots. *Review of Economic Studies*, 75:649–669.
- Bergemann, D. and Morris, S. (2005). Robust mechanism design. *Econometrica*, 73:1771–1813.
- Bierbrauer, F. (2007). Optimal income taxation and public good provision with endogenous interest groups. Mimeo, Max Planck Institute for Research on Collective Goods, Bonn.
- Bierbrauer, F. and Hellwig, M. (2008). Public good provision in a large economy. Work in progress.
- Bierbrauer, F. and Sahm, M. (2008). Optimal democratic mechanisms for income taxation and public good provision. Working Paper, Max Planck Institute for Research on Collective Goods, Bonn.
- Boadway, R. and Keen, M. (1993). Public goods, self-selection and optimal income taxation. *International Economic Review*, 34:463–478.
- Clarke, E. (1971). Multipart pricing of public goods. *Public Choice*, 11:17–33.
- 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.
- Gary-Bobo, R. and Jaaidane, V. (2000). Polling mechanisms and the demand revelation problem. *Journal of Public Economics*, 76:203–238.
- Gaube, T. (2000). When do distortionary taxes reduce the optimal supply of public goods. *Journal of Public Economics*, 76:151–180.
- Green, J. and Laffont, J. (1979). *Incentives in Public Decision-Making*. North-Holland Publishing Company.

- 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.
- Jackson, M. (2001). A crash course in implementation theory. *Social Choice and Welfare*, 18:655–708.
- 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. and Phelan, C. (2007). On the robustness of laissez-faire. Working Paper.
- 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.
- Nava, M., Schroyen, F., and Marchand, M. (1996). Optimal fiscal and public expenditure policy in a two-class economy. *Journal of Public Economics*, 61:119–137.
- Neeman, Z. (2004). The relevance of private information in mechanism design. *Journal of Economic Theory*, 117:55–77.
- Samuelson, P. (1967). Arrow’s mathematical politics. In Hook, S., editor, *Human Values and Economic Policy*. New York University Press.
- Sandmo, A. (1998). Redistribution and the marginal cost of public funds. *Journal of Public Economics*, 70:365–382.
- Stiglitz, J. (1982). Self-selection and pareto-efficient taxation. *Journal of Public Economics*, 17:213–240.

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

A Appendix

A.1 Characterization of robustly implementable social choice functions

An individual is from now on characterized by a preference parameter, a productivity level, beliefs about the the distribution of these characteristics in the economy and beliefs about the beliefs of other individuals. These data are summarized in an *extended type space* $\mathcal{T} := [T, \gamma, \pi]$, where T is the set of extended types, $\gamma : t \mapsto \gamma(t) = (w(t), \theta(t))$, is a mapping that specifies the skill level and the taste parameter for an individual with extended type t . For brevity, I refer to the the pair $\gamma(t) := (w(t), \theta(t))$ as the payoff type of type t and denote the set of possible payoff types by $\Gamma := W \times \Theta$. $\pi : t \mapsto \pi(t)$ specifies the beliefs of an an individual of type t with respect to the cross-section distribution of extended types; i.e., $\pi(t)[X]$ is the probability that this individual assigns to the event “the cross-section distribution of extended types, ϕ , belongs to a set X of probability distributions on T ”.

An extended type space introduces an additional source of heterogeneity among individuals, namely their beliefs about the state s of the economy, but also their higher order beliefs, i.e., their beliefs about the beliefs of other individuals, and so on. Given this enriched description of individuals, I now allow for more sophisticated mechanisms which may depend on the beliefs of individuals.¹⁵

A strategy $\hat{\sigma}$ in the game that is induced by some mechanism $M = (A, Q, C, Y)$ is now viewed as a function that specifies for each extended type t an action $\hat{\sigma}(t) \in A$. $\hat{\sigma}$ is an interim Nash equilibrium for the game induced by M on a given type space \mathcal{T} , if for every t and every a ,

$$\hat{\sigma}(t) \in \operatorname{argmax}_{a \in A} \int_{\Delta_T} U(Q(\rho_{\hat{\sigma}}(\phi)), C(\rho_{\hat{\sigma}}(\phi), a), Y(\rho_{\hat{\sigma}}(\phi), a), \gamma(t)) d\pi(t), \quad (9)$$

where Δ_T denotes the set of probability distributions on T , and $\rho_{\hat{\sigma}}(\phi)$ is the distribution of actions that is induced by $\hat{\sigma}$ if the distribution of extended types is $\phi \in \Delta_T$.

Every type space \mathcal{T} is associated with a cross-section distribution of public goods preferences, $s_{\mathcal{T}}$ and a cross-section distribution of extended types $\phi_{\mathcal{T}}$. A social choice function (q, c, y) is said to be implementable as an interim Nash equilibrium on a given type space \mathcal{T} if there is a mechanism $M = (A, Q, C, Y)$ with an interim Nash equilibrium $\hat{\sigma}$ that satisfies

$$Q(\rho_{\hat{\sigma}}(\phi_{\mathcal{T}})) = q(s_{\mathcal{T}})$$

¹⁵Note that, by definition, a social choice function does not depend on the beliefs of individuals. As in Bergemann and Morris (2005), there must the same economic outcome in two situations where individuals have the same utility functions but differ in their beliefs.

and

$$C(\rho_{\hat{\sigma}}(\phi_{\mathcal{T}}), \hat{\sigma}(t)) = C(s_{\mathcal{T}}, w(t), \theta(t)) \quad \text{and} \quad Y(\rho_{\hat{\sigma}}(\phi_{\mathcal{T}}), \hat{\sigma}(t)) = y(s_{\mathcal{T}}, w(t), \theta(t)),$$

for every t .

Proposition 4 A social choice function is implementable as an interim Nash equilibrium on every type space if and only if it is incentive compatible.

Proof Along similar lines as for Proposition 1 one can show that a social choice function (q, c, y) is implementable as an interim Nash equilibrium by some mechanism M on a given type space \mathcal{T} 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 $Q' : \phi \mapsto Q'(\phi)$, $C' : (\phi, t) \mapsto C'(\phi, t)$, and $Y' : (\phi, t) \mapsto Y'(\phi, t)$ such that

i) Truth-telling is an interim Nash equilibrium; i.e., for all t

$$t \in \operatorname{argmax}_{t' \in T} \int_{\Delta_{\mathcal{T}}} u(Q'(\phi), C'(\phi, t'), Y'(\phi, t'), \gamma(t)) d\pi(t). \quad (10)$$

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

$$Q'(\phi_{\mathcal{T}}) = q(s_{\mathcal{T}}) \quad (11)$$

and

$$C'(\phi_{\mathcal{T}}, t) = c(s_{\mathcal{T}}, t) \quad \text{and} \quad Y'(\phi_{\mathcal{T}}, t) = y(s_{\mathcal{T}}, t), \quad (12)$$

for every t .

We will show that a social choice function is ex post individually incentive compatible if and only if it is truthfully implementable as an interim Nash equilibrium on every extended type space \mathcal{T} .

“ \Leftarrow ”: Given an incentive compatible social choice function (q, c, y) and given an extended types space \mathcal{T} , construct a direct mechanism (T, Q', C', Y') such that (11) and (12) hold. We seek to verify that, for every t ,

$$\begin{aligned} t &\in \operatorname{argmax}_{t' \in T} \int_{\Delta_{\mathcal{T}}} u(Q'(\phi), C'(\phi, t'), Y'(\phi, t'), \gamma(t)) d\pi(t) \\ &= \operatorname{argmax}_{t' \in T} \sum_s \hat{\pi}(t)[s] u(q(s), c(s, \gamma(t')), y(s, \gamma(t')), \gamma(t)), \end{aligned}$$

where $\hat{\pi}(t)[s] := \int_{\{\phi \in \Delta_{\mathcal{T}} | \phi^1 = s\}} d\pi(t)$, and ϕ^1 denotes the marginal cross-section distribution of payoff types associated ϕ . Equivalently, for every t ,

$$\gamma(t) \in \operatorname{argmax}_{\gamma' \in \Gamma} \sum_s \hat{\pi}(t)[s] u(q(s), c(s, \gamma'), y(s, \gamma'), \gamma(t)).$$

This follows from the fact that (q, c, y) is incentive compatible.

“ \implies ”: If (q, c, y) is truthfully implementable as an interim Nash equilibrium on every extended type space, then $[Q, C, Y]$ is, in particular, truthfully implementable on every *complete information type space*; i.e., on every type space \mathcal{T} which is such that, for all t , $\hat{\pi}(t)[s] = 1$, for some given s . If a direct mechanism (T, Q', C', Y') truthfully implements (q, c, y) on such a type space then it must be true that (11) and (12) hold for this type space. Using these conditions to substitute for Q' , C' , and Y' , the equilibrium conditions in (10) becomes: for all t and all t' ,

$$u(q(s), c(s, \gamma(t)), y(s, \gamma(t)), \gamma(t)) \geq u(q(s), c(s, \gamma(t')), y(s, \gamma(t')), \gamma(t)) ;$$

or, equivalently, for all γ and γ' ,

$$u(q(s), c(s, \gamma), y(s, \gamma), \gamma) \geq u(q(s), c(s, \gamma'), y(s, \gamma'), \gamma) .$$

Since the choice of s' was arbitrary, the latter inequality holds for any possible value of s . Hence, (q, c, y) is ex post individually incentive compatible. ■

A.2 Proof of Proposition 1

Suppose the Proposition is false, then there exist $s, (w, \theta)$ and $(\hat{w}, \hat{\theta})$ such that

$$u(q(s), c(s, w, \theta), y(s, w, \theta), w, \theta) < u(q(s), c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \hat{\theta}), w, \theta).$$

Since M implements (q, c, y) this inequality can be equivalently written as

$$\begin{aligned} & u(Q(\alpha^*(s)), C(\alpha^*(s), \sigma^*(w, \theta)), Y(\alpha^*(s), \sigma^*(w, \theta)), w, \theta) \\ & < u(Q(\alpha^*(s)), C(\alpha^*(s), \sigma^*(\hat{w}, \hat{\theta})), Y(\alpha^*(s), \sigma^*(\hat{w}, \hat{\theta})), w, \theta) \end{aligned}$$

This yields a contradiction to the assumption that σ^* is an ex post equilibrium for the game induced by M . ■

A.3 Proof of Proposition 2

To prove the only if-part note that both the inequalities in (7) and the equalities in (6) are implied by incentive compatibility. To prove the if-part, consider a social choice function such that the inequalities in (7) and the equalities in (6) hold, but individual incentive compatibility fails. Then there exist (w, θ) and $(\hat{w}, \hat{\theta})$ and s such that $V(c(s, \theta, w), y(s, \theta, w), w) < V(c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \hat{\theta}), w)$. Using (6) and (7) one has:

$$\begin{aligned} V(c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \hat{\theta}), w) & = V(c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \theta), w) \\ & \leq V(c(s, \theta, w), y(s, \theta, w), w) . \end{aligned}$$

Hence, a contradiction. ■

A.4 Proof of Proposition 3

" \Leftarrow ": Consider a feasible social choice function (q, c, y) . Suppose that a tax function τ satisfying properties i) - iii) exists. Suppose that (q, c, y) is not incentive compatible. Then there exist s , (w, θ) , and $(\hat{w}, \hat{\theta})$ such that

$$u(q(s), c(s, w, \theta), y(s, w, \theta), w, \theta) < u(q(s), c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \hat{\theta}), w, \theta).$$

Using the tax function τ this is equivalent to

$$\begin{aligned} & u(q(s), y(s, w, \theta) - \tau(y(s, w, \theta)), y(s, w, \theta), w, \theta) \\ & < u(q(s), y(s, \hat{w}, \hat{\theta}) - \tau(y(s, \hat{w}, \hat{\theta})), y(s, \hat{w}, \hat{\theta}), w, \theta). \end{aligned}$$

But this contradicts that $\forall s, \forall (w, \theta)$:

$$y(s, w, \theta) \in \arg \max_y u(q(s), y - \tau(s, y), y, \theta, w)$$

" \Rightarrow ": Consider a feasible and incentive compatible social choice function (q, c, y) and construct τ as follows: For any x such that there is s and (w, θ) with $y(s, w, \theta) = x$, define $\tau(s, x)$ by the equation¹⁶

$$\tau(s, x) = y(s, w, \theta) - c(s, w, \theta).$$

Given this function τ , consumption equals after tax income and budget balance holds. For all other levels of x set $\tau(s, x) = x$.¹⁷

Now suppose this function τ does not satisfy the property that $\forall s, \forall (w, \theta)$:

$$y(s, w, \theta) \in \arg \max_y u(q(s), y - \tau(s, y), y, \theta, w)$$

Then there exist (w, θ) , $(\hat{w}, \hat{\theta})$ and s such that

$$\begin{aligned} & u(q(s), y(s, w, \theta) - \tau(y(s, w, \theta)), y(s, w, \theta), w, \theta) \\ & < u(q(s), y(s, \hat{w}, \hat{\theta}) - \tau(y(s, \hat{w}, \hat{\theta})), y(s, \hat{w}, \hat{\theta}), w, \theta), \end{aligned}$$

or, using that for all (w, θ) , $c(s, w, \theta) = y(s, w, \theta) - \tau(s, y(s, w, \theta))$,

$$u(q(s), c(s, w, \theta), y(s, w, \theta), w, \theta) < u(q(s), c(s, \hat{w}, \hat{\theta}), y(s, \hat{w}, \hat{\theta}), w, \theta).$$

This contradicts incentive compatibility. ■

¹⁶Note that his equation uniquely determines $\tau(s, x)$. If not, one would have, for given s , different consumption levels corresponding to the same income requirement; hence, a contradiction to incentive compatibility.

¹⁷It is implicitly assumed that, for any agent, zero consumption implies a utility level of $-\infty$ and that hence the corresponding y is never chosen, whenever there is an alternative with positive consumption available.