**Review
of Economic
Design**

Feasible implementation of taxation methods⋆

Nir Dagan[1], Roberto Serrano[2], Oscar Volij[2]

[1] Department of Economics, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27,
E-08005 Barcelona, Spain
[2] Department of Economics, Brown University, Providence, RI 02912, USA
(e-mail: roberto_serrano@brown.edu)

Abstract. This paper studies implementation of taxation methods in one-commodity environments in which the incomes of the agents are unknown to the planner. Feasibility out of equilibrium imposes that the mechanism depend on the environment. We present two mechanisms. The first one, which requires complete information, implements every taxation method in Nash, strong and coalition-proof equilibrium. The second, where informational requirements are relaxed, implements a large class of consistent and monotone methods in subgame perfect equilibrium. Neither mechanism employs the off-equilibrium devices used by the general theory. Under fully private information no method is implementable.

JEL classification: D31, D71

Key words: Feasible implementation, taxation methods, consistency, decentralization, flat tax

1. Introduction

Already since the Classical school, the topic of taxation has drawn constant attention from economists. Two early inquiries are found in Ricardo (1817) and Mill (1848). For the most part, authors have agreed that reasonable taxation schemes should rely on two fundamental principles: efficiency and equity. Accordingly, two theories of taxation have been developed, each emphasizing one of these two properties.

The modern theory of optimum taxation was greatly developed in the seventies, due to the influential work of Mirrlees (1971; see also 1986 for a survey). In economies where agents are making consumption/labor decisions, the question asked by this literature is to what extent the introduction of an income tax

---

distorts the allocation of resources. Its findings are mainly negative: second-best allocations usually arise. Within this approach, it is assumed that the government's goals are described by a social welfare function which depends only on the after-tax allocation, thereby disregarding property rights altogether.

The equity approach, adopted by Bentham (1789), Mill (1848), Edgeworth (1898), Young (1987) and others, and under which this paper falls, concentrates on normative issues. It poses questions that relate income distribution and property rights. Agents' incomes are given exogenously, which means that the labor (production) decision has already taken place. In addition, the level of taxes that the tax authority wants to collect (perhaps in order to finance an efficient level of public expenditure) has also been decided. The problem is then how to split the amount of taxes among the agents, taking into account their taxable incomes.[1]

Recently, the machinery of axiomatic game theory has been successfully applied to this approach. Young (1988) characterizes Mill's equal sacrifice methods and Dagan (1994) axiomatizes Bentham's least sacrifice method. Both characterizations use the consistency property (see Sect. 2 below and Thomson 1990). These axiomatic criteria provide us with answers to the question of how to divide "fairly" the tax burden among the agents in society. In this paper, we bring incentives to the forefront, without imposing a specific normative criterion. That is, if the tax authority does not know the agents' incomes and it wishes to implement a certain taxation method, can the truth be elicited so that the "right" tax allocation results?

Our answer to this question will be affirmative (if information is not fully private) because we do not require dominant strategy implementability (that is, that every agent tell the truth regardless of the others' announcements). This strong requirement may explain Mirrlees's negative results: by using tax schedules in which an agent's tax depends only on his income, the actions of the others cannot affect an agent.[2] We are able to implement any taxation method in Nash, strong and coalition-proof equilibrium, and a large class of consistent taxation methods (that contains those of equal sacrifice) in subgame perfect equilibrium.

We shall assume that the tax authority has incomplete information about the true vector of taxable incomes, which (in Sect. 4) is common knowledge among all agents of society. The latter is not the most reasonable assumption when studying income taxation in a large economy. Thus, one interpretation of our exercise in Sect. 4 is that these agents are members of a club that wishes to collect monetary contributions from them (the countries of the European Union, for example).

---

[1] One may argue that this approach misses the big picture by disregarding the issue of efficiency. One possible interpretation, though, is that agents in the economy have quasilinear preferences over the level of public expenditure. Thus, its efficient level is independent of income distribution.

[2] In an economy a la Mirrlees, Piketty (1993) also obtains positive results of implementation of first-best allocations using Bayesian equilibrium (it is an equilibrium to report the truth). This is done via generalized tax schedules, in which an agent's tax depends on the entire vector of incomes of the economy. We explain in our concluding remarks that implementation in dominant strategies is not possible in our framework.

The mechanism in Sect. 5 may be adapted to environments where the informational requirements are weaker. This is done in Sect. 6. The property of consistency in taxation has been defended as a coherent way to generalize a bilateral principle of justice to many agent contexts. In addition, our Sects. 5 and 6 show that consistency of a taxation method may help to decentralize its implementation, as long as there exists an "informationally big" player, i.e., a player who has complete information. This assumption can be relaxed if the planner wishes to implement the flat (proportional) tax method. Then, the implementation can be decentralized through certain small groups of agents (firms, town districts, etc.), where the existence of an "informationally big" player in each of them is the key.

In contrast with our previous positive results, Sect. 7 proves an impossibility theorem. If the economy is one with fully private information, no taxation method can be implemented. Agents always have an incentive to under- report their incomes. It is interesting to compare our two extreme results in our framework. When information is complete among the agents, every taxation method can be implemented, while none if information is private. The explanation lies on the agents' possibility of policing each other in the misrepresentation of one's income. This suggests that the tax authority should make use of all possible information flows in the economy, either by making agents monitor one another or by instituting efficient auditing policies. Our mechanisms allow the tax authority to check that agents do not over-report incomes, but nothing beyond that. That is, as opposed to the auditing and tax evasion literature, we assume that it is not possible to discover under- reporting.

Our framework is one in which preferences are known to the designer (utility is strictly increasing in money), while the feasible set (the vector of taxable incomes) is not. This departs significantly from most work on implementation theory, which has concentrated on the opposite case, including both the extensive work that originates in Maskin (1977) and the papers on taxation mentioned above.[3] In particular, the notion of a *mechanism* as a mapping from message profiles to feasible allocations is not straightforward, as the feasible set itself is an unknown to the designer. Our first precursor is the work on *feasible implementation* by Hurwicz et al. (1995). Due to the unknown feasible set, the added difficulty that the problem at hand presents is the issue of feasibility out of equilibrium. We follow Hurwicz et al. (1995) in order to deal with this problem: in particular, we show in Sect. 3 that feasible implementation imposes that the mechanism must depend in a non-trivial way on the true state of the world. Other papers that deal with feasible implementation are Postlewaite and Wettstein (1989) for the Walrasian correspondence, Tian (1989,1993) for the Lindahl correspondence, Serrano and Vohra (1997) for the Core and Tian (1994) for the Linear Cost Share Equilibrium correspondence.

In his 1992 survey, Moore suggests that implementation theory should turn to applications and stress the realism of its mechanisms. As suggested there, when

---

[3] See Maskin (1985) and Moore (1992) for detailed surveys of the subject.

dealing with a particular application, one should expect "nicer" mechanisms than the abstract ones used by the general theory. Our mechanisms for taxation fit very well Moore's goals, by dispensing with devices such as integer and modulo games, non-compact strategy sets, large off-equilibrium punishments, and so on (see also Jackson (1992) for a critique of these devices). The two mechanisms we present are simple enough: in one of them, every agent reports all incomes to a central agency; in the other the designer chooses a tax collector, who makes a tax proposal, and agents bargain over taxes.

2. The taxation model

We present a taxation problem, first introduced by O'Neill (1982) and Aumann and Maschler (1985) in the context of bankruptcy. An *economy* or *taxation problem* is a pair $(\mathbf{x}, T)$, where $\mathbf{x} \in \mathbb{R}_+^N$ is the list of taxable incomes, N is a finite set of agents that contains at least two members, and $0 \leq T \leq \sum_{i \in N} x_i$ is the total amount to be collected. A *tax allocation* is a list $\mathbf{t} \in \mathbb{R}^N$ such that $\mathbf{0} \leq \mathbf{t} \leq \mathbf{x}$, $\sum_{i \in N} t_i = T$, that is, subsidies to agents are ruled out and a total amount of taxes is to be collected (for any two vectors $\mathbf{a}, \mathbf{b}$, we write $\mathbf{a} \leq \mathbf{b}$ to mean $a_i \leq b_i$ for all i). We denote the set of all tax allocations of $(\mathbf{x}, T)$ by $A(\mathbf{x}, T)$.

A *taxation method* is a function f that assigns a tax allocation to each economy. A taxation method imposes taxes upon the agents as to collect the amount T. Note that taxes are not employed to conduct transfer payments among the agents in the economy.

Examples

a) Equal sacrifice methods Let $U : \mathbb{R}_{++} \to \mathbb{R}$ be a continuous and strictly increasing function that satisfies $lim_{y \to 0} U(y) = -\infty$. The equal sacrifice method f relative to U satisfies

$$f(\mathbf{x}, T) = \mathbf{t} \leftrightarrow \exists c \geq 0 \text{ such that } \forall i \in N \text{ with } x_i > 0, U(x_i) - U(x_i - t_i) = c.$$

These methods assign taxes so as to equalize absolute sacrifice evaluated according to a prespecified utility function. The equal sacrifice method with respect to the logarithmic function is the proportional taxation method or flat tax. The equal sacrifice principle in taxation appears in Mill (1848, Book V).

b) The head tax: $t_i = \min\{\lambda, x_i\}$;

c) The levelling tax: $t_i = \max\{x_i - \lambda, 0\}$;

d) The Cohen-Stuart (1889) - Edgeworth (1898) family of progressive methods: $t_i = \max\{0, x_i - x_i^\lambda\}$;

e) The Cassel (1901) methods: $t_i = x_i^2/(x_i + \lambda)$; where λ is determined so that $\sum_{i \in N} t_i = T$.

A method f is *T-monotone* or simply monotone if for all $(\mathbf{x}, T)$ and $0 \leq T' \leq T$, $f(\mathbf{x}, T') \leq f(\mathbf{x}, T)$. Monotonicity says that a decrease in the amount of taxes

to be collected does not harm any agent. A method f is *strictly monotone* if for all $(\mathbf{x}, T)$ and $0 \le T' < T$, if $x_i > 0$ then $f_i(x, T') < f_i(x, T)$.

A method f is *consistent* if for any finite non-empty set I of agents, for all $(\mathbf{x}, T)$, $\mathbf{x} \in \mathbb{R}_+^N$, and for all $\emptyset \ne J \subset I$:

$$\mathbf{t} = f(\mathbf{x}, T) \Rightarrow \mathbf{t}|\mathbb{J} = f(\mathbf{x}|J, \sum_{i \in Jt_i}),$$

where if $\mathbf{y} \in \mathbb{R}_+^I$, $\mathbf{y}|J$ is the projection of $\mathbf{y}$ on $\mathbb{R}_+^I$. Suppose that a method f assigns the tax allocation $\mathbf{t}$ to the economy $(\mathbf{x}, T)$. Suppose also that some subset of agents wants to reallocate the total amount of taxes $\sum_{i \in J} t_i$ assigned to them. If we apply the same method f to allocate this tax burden among these agents, each of them will get the amount originally assigned to him, provided f is consistent.

Consistency and strict monotonicity were the key properties used by Young (1988) to axiomatize the equal sacrifice methods. All the methods mentioned above are consistent and monotone, although only the Cassel and equal sacrifice methods are strictly monotone.

3. The implementation problem

We consider the situation where the mechanism designer (the tax authority) knows the set of agents N and the amount T to be collected, but does not know the agents' incomes $\mathbf{x}$. However, the designer knows that the amount T can be collected from the agents. On the other hand, the economy $(\mathbf{x}, T)$ is common knowledge among all agents. In the language of implementation theory, the set of taxation environments for a given set of agents N and a budget T is

$$Z(N, T) := \{(\mathbf{x}, T) | \mathbf{x} \in \mathbb{R}_+^n, 0 \le T \le \sum_{i \in N} x_i\}.$$

In the economies we consider, there is a single commodity – money. Further, each agent's preference is dependent only on his own after-tax income and it is strictly increasing in this variable. As this is known by the tax authority (who, on the other hand, does not know the agents' incomes), we depart from most of the literature on implementation of social choice correspondences, which has concentrated on implementation problems with unknown preferences and known feasible set.

Let Z be a non-empty class of economies. A mechanism for Z is a function G that assigns to each economy $(\mathbf{x}, T) \in Z(N, T)$ a game form $G(\mathbf{x}, T)$ defined as follows: $G(\mathbf{x}, T) \equiv \langle (M_i)_{i \in N}, g \rangle$, where M_i is the set of messages (strategies) for player i, and $g : \prod_{i \in N} M_i \to A(\mathbf{x}, T)$ is an outcome function that assigns a tax allocation to each n-tuple combination of messages. A mechanism $G(Z)$ *implements* in a certain equilibrium concept the taxation method f on the class of environments $Z(N, T)$ if for all economies $(\mathbf{x}, T) \in Z(N, T)$, all the equilibrium outcomes of $G(\mathbf{x}, T)$ are $f(\mathbf{x}, T)$.

Our definition of a mechanism for $Z(N,T)$ does not preclude the possibility that G be a constant function. Proposition 3.1 below, however, shows that such a mechanism cannot implement any taxation method on any interesting class of economies.

Proposition 3.1. *Consider the class of environments $Z(N,T)$. Let f be a taxation method. Any mechanism G that implements f is dependent on the environments. That is, for all two economies $(\mathbf{x},T)$ and $(\mathbf{x}',T) \in Z(N,T)$, if $f(\mathbf{x},T) \neq f(\mathbf{x}',T)$, then $G(\mathbf{x},T) \neq G(\mathbf{x}',T)$.*

Proof. Let G be a mechanism that implements f on $Z(N,T)$. Let $(\mathbf{x},T)$ and $(\mathbf{x}',T)$ be two economies in $Z(N,T)$ with $f(\mathbf{x},T) \neq f(\mathbf{x}',T)$. Assume by way of contradiction that $G(\mathbf{x},T) = G(\mathbf{x}',T)$.

Because G implements f, there exist equilibrium messages $\mathbf{m}^*$ such that $g(\mathbf{m}^*) = f(\mathbf{x},T)$. Furthermore, since $G(\mathbf{x},t) = G(\mathbf{x}',t)$, $f(\mathbf{x},T)$ is also an equilibrium outcome of $G(\mathbf{x}',t)$, which is a contradiction. $\square$

Remark: Proposition 3.1 is independent of the equilibrium concept considered.

Hurwicz et al. (1995) proved a similar result in the context of implementation in exchange economies in which the planner knows the agents' preferences but does not know the endowments. A necessary condition for implementability in our context is thus the dependence of the game form on the vector $\mathbf{x}$ of taxable incomes. Of course, this cannot mean that the designer knows the environment beforehand; likewise, we should not assume that the feasible set of incomes can be fully verified, as would be the case if the planner could confiscate every agent's income. In either of these two cases, implementation becomes trivial: a mechanism is simply not needed since it is possible to impose directly the desired taxation method. However, in view of Proposition 3.1 and given the requirement that the mechanism assign feasible allocations in and out of equilibrium, partial verifiability of the feasible set is required.

We shall therefore assume, like Hurwicz et al. (1995), that the planner can verify the announced incomes: every player must report incomes to the planner, and show the reported amount. That is, while reporting less than one's income is possible, exaggeration is ruled out.

Formally, we say that a mechanism $G(Z)$ is *acceptable* if for all economies $(\mathbf{x},T) \in Z(N,T)$ and for all $i \in N$ player i's strategy set M_i in $G(\mathbf{x},T)$ is of the form $X \times I_i$ where $X = (\prod_{i \in N} X_i) \cap Y$, $X_i = [0,x_i]$, and I_i and $Y \subset \mathbb{R}^N$ are non-empty sets independent of $\mathbf{x}$. Acceptability requires that the mechanism G depend on the environment only through the set X, which by definition does not allow any agent to overstate anyone's income. As for Y and I_i, they are independent of the environments. For example, Y may be the set of $\mathbf{z}$ such that $(\mathbf{z},T)$ is an economy, whereas I_i is agent-specific.

Proposition 3.1 also clarifies the issue that the words "mechanism" and "game form," that are generally used interchangeably, refer to different objects in this context. The former refers to the class of "game forms" that is needed to perform the implementation over the class of environments $Z(N,T)$, while the latter refers

to the game for a particular economy. From now on, we shall abuse language slightly and talk about a "mechanism" even when we refer to the "game form" for a particular economy.

4. Normal form implementation

In this section we construct an acceptable canonical mechanism in normal form that implements every taxation method. In contrast to the canonical mechanisms of the theory of implementation, our game form is valid in environments with *either two* or more than two agents.

Let $N = 1, 2, \ldots, n$ with $n \geq 2$. Define the mechanism $\beta(Z(N, T))$ as follows: for each economy $(\mathbf{x}, T)$ and for all $i \in N$, a typical message m_i is a list $\mathbf{y}^i = (y_1^i, y_2^i, \ldots, y_n^i)$ of reported incomes in which $\mathbf{0} \leq \mathbf{y}^i \leq \mathbf{x}$ and $\sum_{i \in N} y_i^j \geq T$. Denote by $\mathbf{y}$ the componentwise maximum of the agents' reports, i.e., $y_i = \max\{y_i^j | j \in N\}$. For a message profile $\mathbf{m}$, define the outcome function g as $g(\mathbf{m}) = f(\mathbf{y}, T)$.

In designing the sets of strategies, we make heavy use of the complete information assumption. The vectors $\mathbf{y}^i$ must be verified. Each agent polices the others and acts as a confidant to the designer, providing him with information about the entire profile of incomes. As shown below, notice that not even coalitional arrangements can perturb the "efficient" mutual policing.

Theorem 4.1. *Let f be a taxation method. The mechanism β implements f in Nash, strong Nash and coalition proof Nash equilibria over the class $Z(N, T)$.*

Proof. The unanimous report of $\mathbf{x}$ is a Nash equilibrium. Any individual deviation does not change the componentwise maximum incomes announced and thus the same tax allocation results.

There cannot be an equilibrium with an outcome other than $f(\mathbf{x}, T)$. Assume by way of contradiction there is such an equilibrium. Then there is a player i that pays a higher tax than he would have paid if the true incomes were reported. Hence, this player has a profitable deviation: to announce the true incomes and pay $f_i(\mathbf{x}, T)$.

Note finally that the Nash equilibrium suggested above is also a strong Nash. The only coalition that can change the tax allocation is the coalition of all agents. However, since the total tax levied must be T, it cannot be that all agents are better off by deviating. Finally, it is also a coalition proof Nash equilibrium as any strong Nash is coalition proof.[4]   □

Remarks: *I)* Since the mechanism maps messages into tax allocations, it is not wasteful: the tax authority always collects T (both in and out of equilibrium).
II) No properties of the taxation method f are required in Theorem 4.1. Note that the deviation of the agent used in the proof does not say that he maximizes

---

[4] Strong Nash equilibrium requires the strategies to be immune to all coalitional deviations. Coalition proof Nash equilibrium considers only "self-enforcing" coalitional deviations. For formal definitions, see Bernheim et al. (1987).

the other agents' reported incomes, but announces the true incomes which may include a higher report of his own income. This argument holds independent of any properties of the taxation method. Since we do not assume that the taxation method is increasing in an individual's income, maximizing the other agents' incomes in the report is not necessarily a dominant strategy.

5. Extensive form implementation of consistent methods

Despite its many advantages, an objection against the mechanism β in a private ownership economy is that it is not "privacy-preserving." That is, every agent is responsible for the report to the tax authority of the entire profile of incomes. It would be more realistic to think that agents "mind their own business," and that they resort to comparisons of incomes before the tax authority only when they wish to challenge a given allocation of taxes. This is a virtue of the mechanism presented in this section.

The next mechanism is an extensive form that implements a large class of consistent and monotone methods in subgame perfect equilibrium. We restrict our analysis to the case where all incomes are positive. That is, denote by $Z^*(N,T)$ the following subclass of $Z(N,T)$:

$$Z^*(N,T) := \{(\mathbf{x},T) \in Z(N,T) | \mathbf{x} \in \mathbb{R}_{++}^N\}.$$

The following mechanism will allow us in Sect. 6 to dispense with the assumption of a complete information environment. Although stronger than necessary, we maintain it in this section.

Let $((x_1, x_2, ..., x_n), T) \in Z^*(N,T)$ be an n-person economy. We describe informally the game $\gamma^f(\mathbf{x}, T)$ as follows. The planner chooses player 1 to be its tax collector. Player 1 must make a tax proposal to which the rest of the players must respond. If a player i accepts the proposal, player 1 has the power to enforce that player i will pay those taxes. If player i rejects, he renegotiates his taxes with the tax collector using the game of Sect. 4. These bilateral renegotiations determine the tax paid by each rejector. Finally, the tax collector pays the remainder of the tax bill.

Formally, player 1 (the agent with index 1) proposes a tax allocation $\mathbf{t}$ in $A(\mathbf{x}, T)$. Following this proposal all the other agents respond sequentially, either by accepting or rejecting the offer. The order of responses follow the protocol induced by the agents' indices, namely agent 2 is the first to respond and agent n is the last one. In order to continue with the definition of the game and to define the players' payoffs, we need to define an auxiliary variable of "interim tax payments" for the proposer. Given a proposal $\mathbf{t}$, define $w_1^1 = t_1$. Assume now that w_1^{i-1} has already been defined for all k between 2 and $i(\leq n)$. If player i accepts the proposal $\mathbf{t}$, his tax payment is t_i. In this case set $w_1^i = w_1^{i-1}$ and it is $i + 1$'s turn to respond, unless $i = n$ in which case the game ends. If i rejects, player 1 and i conduct renegotiations using the game form $\beta[(x_1, x_i); w_1^{i-1} + t_i]$ to determine i's payoff and w_1^i (it is easy to see that $[(x_1, x_i); w_1^i - 1 + t_i]$ is

indeed an economy). The amount w_1^i is determined to be 1's tax payment in the bilateral negotiation and i's tax payment is his own share. After all players have responded, player 1 pays w_1^n and the game ends.

Theorem 5.1. *Let f be a consistent and strictly monotone taxation method. The mechanism γ^f implements f in subgame perfect equilibrium on $Z^*(N, T)$.*

Proof. Fix an n-player economy $(\mathbf{x}, T) \in Z^*(N, T)$ and consider the game $\gamma^f(\mathbf{x}, T)$. Let $\mathbf{t}^* = f(\mathbf{x}, T)$. The proof that there is a unique subgame perfect equilibrium payoff of γ^f is organized in five lemmas.

Lemma 5.1.1. *In any subgame perfect equilibrium of $\gamma^f(\mathbf{x}, T)$ player 1's equilibrium tax payment is at most t_1^*.*

Proof. Player 1 can guarantee that his taxes are at most t_1^* by proposing $\mathbf{t}^*$ and reporting the truth in the bilateral negotiations. By adopting this strategy, in any bilateral renegotiation $w_1^i = f_1[(x_1, x_i); w_1^{i-1} + t_i^*]$. By consistency of f and a straightforward induction argument, this last expression equals t_1^* for all i.   □

Lemma 5.1.2. *Let s be a subgame perfect equilibrium of $\gamma^f(\mathbf{x}, T)$ and consider a node in the game tree where player $i > 1$ has to respond to proposal $\mathbf{t}$. Let w_1^{i-1} be player 1's current interim tax and let z_i be i's equilibrium payoff in this subgame. Then, according to σ, the subsequent interim tax w_1^i will be equal to $\max\{w_1^{i-1}; f_1[(x_1, x_i); w_1^{i-1} + t_i]\}$ and $z_i = \min\{t_i; f_i[(x_1, x_i); w_1^{i-1} + t_i]\}$.*

Proof. The proof is by backwards induction. Let $i = n$. If player n accepts, he pays t_n and player 1 pays $w_1^{i-1} = w_1^{n-1}$. If player n rejects the proposal $\mathbf{t}$, they go to the renegotiation game $\beta[(x_1, x_n); w_1 n - 1 + t_n]$ after which the game ends. By Theorem 4.1, the unique equilibrium outcome of this renegotiation game is $f[(x_1, x_n); w_1^{n-1} + t_n]$. Therefore, since σ is a subgame perfect equilibrium we must have $z_n = \min\{t_n; f_n[(x_1, x_n); w_1^{n-1} + t_n]\}$ and $z_1 = w_1^n = \max\{w_1^{n-1}; f_1[(x_1, x_n); w_1^{n-1} + t_n]\}$. Note that since f is strictly monotone, player 1's tax payment $z_1 = w_1^n$ is strictly increasing in w_1^{n-1}.

Assume now that the statement of the lemma is true for $i = n, \ldots, k + 1$ and let $i = k$. If k accepts, he will pay t_k. If he rejects, 1 and k go to the bilateral renegotiation game $\beta[(x_1, x_k); w_1^{k-1} + t_k]$. In this stage some w_1^k will be determined in equilibrium. By the induction hypothesis, 1's final tax payment is strictly increasing in w_1^k and by the rules of the game, k's payment is strictly decreasing in w_1^k. Therefore 1 and k face a strictly competitive situation. Moreover, both players can guarantee, by reporting the truth, that $w_1^k = f_1[(x_1, x_k); w_1^{k-1} + t_k]$, which means that in equilibrium this is going to be the interim tax payment. As a result we have that, since σ is a subgame perfect equilibrium,

$$z_k = \min\{t_k; f_k[(x_1, x_k); w_1^{k-1} + t_k]\} \text{ and } w_1^k = \max\{w_1^{k-1}; f_1[(x_1, x_k); w_1^{k-1} + t_k]\}$$

□

Lemma 5.1.3. *Let $\mathbf{z}$ be an equilibrium outcome of $\gamma^f(\mathbf{x}, T)$. Then $f_1[(x_1, x_i); z_1 + z_i] \leq z_1$ for all i between 2 and n.*

Proof. Let σ be an equilibrium outcome of $\gamma^f(\mathbf{x}, T)$ and let $\{w_1^i\}_{i=1}^n$ be the interim taxes of player 1, determined by the equilibrium path. According to Lemma 5.1.2,

$$*. \qquad\qquad z_i = \min\{t_i; f_i[(x_1, x_i); w_1^{i-1} + t_i]\}$$

Moreover, in equilibrium,

$$**. \qquad\qquad w_1^i = w_1^{i-1} + t_i - z_i$$

By $(**)$ and the definition of z_1, we have

$$***. \qquad\qquad z_1 = w_1^n \geq w_1^i \text{ for all } i$$

Given the identity $f_1[(x_1, x_i); z_1 + z_i] + f_i[(x_1, x_i); z_1 + z_i] = z_1 + z_i$, it follows from $(***)$ and monotonicity of f that $f_1[(x_1, x_i); z_1 + z_i] + f_i[(x_1, x_i); w_1^i + z_i] \leq z_1 + z_i$. Using $(**)$ we have $f_1[(x_1, x_i); z_1 + z_i] + f_i[(x_1, x_i); w_1^{i-1} + t_i] \leq z_1 + z_i$ and by $(*)$, $f_1[(x_1, x_i); z_1 + z_i] \leq z_1$. $\qquad\square$

Lemma 5.1.4. *Let $\mathbf{z}$ be an equilibrium outcome of $\gamma^f(\mathbf{x}, T)$. Then $z_1 \geq t_1^*$.*

Proof. If $z_1 < t_1^*$, since $\mathbf{z}$ and $\mathbf{t}^*$ are both tax allocations, there must be some agent i with $z_i > t_1^*$. There are two cases.
Case 1: $z_1 + z_i \geq t_1^* + t_i^*$.
In this case, by monotonicity and consistency, $f_1[(x_1, x_i); z_1 + z_i] \geq f_1[(x_1, x_i); t_1^* + t_i^*] = t_1^* > z_1$.
Case 2: $z_1 + z_i < t_1^* + t_i^*$.
In this case, by monotonicity and consistency, $f_i[(x_1, x_i); z_1 + z_i] \leq f_i[(x_1, x_i); t_1^* + t_i^*] = t_i^* < z_i$.
Therefore, by efficiency, $f_1[(x_1, x_i); z_1 + z_i] > z_1$.

But then in both cases we have $f_1[(x_1, x_i); z_1 + z_i] > z_1$ which contradicts Lemma 5.1.3 $\qquad\square$

Lemma 5.1.5. *Let $\mathbf{z}$ be an equilibrium outcome of $\gamma^f(\mathbf{x}, T)$. Then $\mathbf{z} = \mathbf{t}^*$.*

Proof. Assume that $\mathbf{z} \neq \mathbf{t}^*$. Then, there is a player i, with $z_i > t_i^*$. By Lemmas 5.1.1 and 5.1.4 we know that $t_1^* = z_1$. Therefore this player cannot be the proposer, and we have $z_1 + z_i > t_1^* + t_i^*$. By strict monotonicity of f, $f_1[(x_1, x_i); z_1 + z_i] > f_1[(x_1, x_i); t_1^* + t_i^*]$, and by lemma 5.1.3 we have $z_1 \geq f_1[(x_1, x_i); z_1 + z_i]$. But then, it must be that $t_1^* > f_1[(x_1, x_i); t_1^* + t_i^*]$ which contradicts consistency of f. $\qquad\square$

As for the existence part, it can be checked that the following n-tuple of strategies constitutes a SPE: player 1 offers $\mathbf{t}^*$ and reports the truth in every bilateral renegotiation; each responder i accepts $\mathbf{t}$ if and only if $t_i \leq f_i[(x_1, x_i); w_1^{i-1} + t_i]$ and reports the truth in every bilateral renegotiation.

Therefore, the mechanism γ^f implements f in SPE on $Z^*(N, T)$. $\qquad\square$

Remark: The mechanism is non-wasteful. Also, if the taxation method is anonymous, so is the mechanism, in the sense that it is invariant to any renaming of

the agents. The mechanism imposes the natural protocol just for convenience. The naming of the players can be any other.

An interesting class of taxation methods satisfies both consistency and monotonicity, but not strict monotonicity (e.g., the head tax, the levelling tax, the Cohen Stuart/Edgeworth methods). However, all these methods share the property that, in the tax allocation they prescribe, the share of the richest player is strictly monotone in the total tax T. For these methods we can make the following observation.

Remark: Suppose $x_1 \geq x_i$ for all $i \in N$. Let f be a consistent and monotone taxation method such that player 1's tax payment is strictly monotone in T. Then, the mechanism γ^f implements f in subgame perfect equilibrium on $Z^*(N, T)$. To see this, note that in the proof of Theorem 5.1, strict monotonicity is needed only for the proposer's share.

For these methods the mechanisms γ^f when the proposer is not one of the richest agents in the economy, may yield a continuum of equilibrium payoffs, as shown in Example 1. Therefore, to implement them with this game form, the designer would require an additional piece of information: the player with the highest income (and no other) must be the tax collector.

Example 1: Let $(\mathbf{x}, T) = ((2{,}000, 12{,}000, 1{,}000), 7{,}000)$ and let f be the head tax method. Then, $f(\mathbf{x}, T) = (2{,}000, 4{,}000, 1{,}000)$. Then, the reader can check that the set of SPE tax payments of $\gamma^f(\mathbf{x}, T)$ is $\{(2{,}000, 4{,}000 + e, 1{,}000 - e)$ for $0 \leq e \leq 1{,}000\}$.

6. Some departures from complete information

Thus far the analysis of Sects. 4 and 5 has assumed that the environment was one of complete information. While this is a plausible assumption if our problem is to raise monetary contributions from the members of a small club, it is clearly inappropriate for a taxation problem in a large economy. However, notice that the informational requirements of the consistency-based mechanism γ^f in Sect. 5 can be substantially relaxed as long as there exists one "informationally big" player. Specifically, while the proposer must have complete information about the n players' incomes, the $n - 1$ responders need only to know the proposer's income (besides his own, of course). This requires the game to satisfy that the "interim" outcomes achieved in the negotiations under β are publicly known.

This suggests the following scenario. In a centrally planned economy where every agent works for the center, tax collection could be left to the center itself. Using the mechanism γ^f of Sect. 5 the center sends a tax proposal to each of the workers, who can either accept it or renegotiate their tax payment before the tax authority. The necessary assumption is that the center knows each agent's income and each agent knows the center's "profits." If the taxation problem is the collection of contributions in a small club, the elected president (who may have complete information through access to the club's files) should make the tax proposal.

One shortcoming of the theory we present is the need of the "information-ally big" player. One may wonder about how much can be said in regards to implementation under less stringent informational assumptions. Unfortunately, this does not seem to yield positive results. For example, suppose that the n agents are arranged in a circle and that each agent knows only his income and that of his two neighbors. Then, a mechanism that pins down each pair of neighbors using consistency need not yield the desired tax allocation for the entire economy. Suppose that the taxation method is the head tax. Let the taxation problem be $(\mathbf{x}, T) = ((2{,}000, 500, 12{,}000, 500, 1{,}000, 500), 8{,}500)$. Notice that $f(\mathbf{x}, T) = (2{,}000, 500, 4{,}000, 500, 1{,}000, 500)$. However, the tax allocation $\mathbf{t} = (2{,}000, 500, 4{,}500, 500, 500, 500)$ satisfies that for all neighbors i and j $(t_i, t_j) = f((x_i + x_j), t_i + t_j)$.

In the case of the flat tax, on the other hand, the informational requirements are smaller since we can decentralize the implementation across many small groups. Then, we shall require that certain aggregate information is available and that there is one "informationally big" player in each smaller group.

For instance, suppose the U.S. Federal Government wishes to implement the flat tax method f. The set of agents N is then a very large finite set. The Federal Government may know the amount T it wants to collect and it has data on the GNP $\sum_{i \in N} x_i$, so that it knows that $\sum_{i \in N} x_i \geq T$. In addition, the Federal Government may possess data on the aggregate incomes of each of the states, that is, partitioning the set N into $N = \{N(1), \ldots, N(50)\}$, it knows the vector $(x^{N(1)}, \ldots, x^{N(50)})$ of aggregate state incomes. The Federal Government can then delegate the tax collection to the different states, expecting from each state $N(i), i = 1, \ldots, 50$, an amount $T^{N(i)}$ such that $f((x^{N(1)}, \ldots, x^{N(50)}), T) = (T^{N(1)}, \ldots, T^{N(50)})$.

Each of the states may have information about the aggregate incomes of each of its counties and therefore can delegate the tax collection of the amount $T^{N(i)}$ on them; similarly, each county can do the same with the towns, each town, with its districts. Say this is the last step of decentralization. We can assume that the district representative, who has been elected to office, knows the people in the district well enough to have complete information about their incomes (or he/she can have access to those data). On the other hand, the tax payers in the district also know his/her income (after all, he/she was elected by them and at the time they gathered enough information about the candidate). Then, the mechanism γ^f can be applied to each of the districts and the proposer must be the elected officer.

As reported in Adams (1993), this procedure resembles the one that Augustus introduced in the Roman Empire: "To develop a more uniform system, Augustus decreed that the entire empire should be appraised for tax reform. A great census should be taken to register everybody and everything ... The registration ordered by Augustus did not mean that taxes were to be controlled by the central government. To the contrary, the census involved a decentralization of the tax system. Individual tax payers would never again face Roman tax men. Tax collection would be a local matter administered by local people. The census let the

Roman government know how much wealth and how many people were in the empire. With this information, Rome could make a more equitable assessment against every city and leave collection methods for local administration. Previously, Roman publicani had created a direct link from tax payers to the central government. The new system of Augustus took the Roman government out of the process of assessing and collecting taxes on an individual level" (p. 94).

Other examples come to mind, where a corporate flat tax that must be collected by the central authority can be decentralized across different productive sectors, in which the largest firm in a given sector (who plausibly may have an informational advantage) must act as the tax collector. Alternatively, imagine an income/corporate flat tax where corporations are considered like other individuals (their profits are their taxable incomes). In this case, each firm may be the tax collector for the central tax authority: each firm knows the wages of all its employees through payroll, while the employees know the firm's profits.

Information in this economy is very much decentralized. In particular, the small agents need not know anything about each other's income, while the tax collector in each group needs information only about the group. Also, the central tax authority needs only the aggregate intermediate data to delegate its functions. If the taxation method is other than the flat tax, this scheme will not work in general: additional information beyond the sum of agents' incomes in a group is needed to obtain the desired tax allocation. For example, suppose we are interested in the levelling tax method, there are 10 agents who live in two regions (2 in the north, 8 in the south), each of them has an income of $1,000 and we want to collect $1,000. Then we would like to collect $100 from each agent. However, the decentralization scheme through regions would give us that the agents in the north should pay nothing while each agent in the south should pay $125 : $f((2,000, 8,000), 1,000) = (0, 1,000)$.

7. Fully private information

The discussion in the previous section should be suggestive that implementation becomes more difficult in our environments the more private information is there in the economy. This section confirms this intuition with a negative result in the case of fully private information.

Suppose that each agent $i \in N$ knows only his income x_i and does not know the incomes of any other agent. That is, agent i has a non-degenerate conditional density function $P_i(\mathbf{x}_{-i}|x_i)$ over the incomes $\mathbf{x}_{-i}$ of all other agents. This density assigns positive probability only to sets of values for $\mathbf{x}_{-i}$ such that $((0, \mathbf{x}_{-i}), T)$ is an economy. Denote this class of environments by $Z'(N, T)$. This is consistent with the assumption of fully private information, especially plausible in large economies.

We model this as a problem of incomplete information and the game theoretic solution concept we use is Baysian equilibrium: each possible type of each agent i must play a best response to the other agents, given the information he holds about them, that is, given $P_i(\mathbf{x}_{-i}|x_i)$. The following is our next result:

Theorem 7.1. *Let f be a taxation method. The method f is not incentive compatible over the class $Z'(N, T)$. Therefore, there does not exist a mechanism G that implements f in Bayesian equilibrium over this class.*

Proof. By the revelation principle, there is no loss of generality in restricting attention to direct mechanisms, that is, those in which each agent is asked to report his income. We shall prove that there does not exist a direct mechanism G with a Bayesian equilibrium that yields taxes equal to $f(\mathbf{x}, T)$ for all possible $(\mathbf{x}, T)$ in the class.

Suppose not. That is, assume that G is a direct mechanism with such a Bayesian equilibrium. We will show that agent i with income $x_i > 0$, who pays positive expected taxes, has a profitable deviation: to imitate another type of agent i with income $x_i' = E[f_i((x_i, \mathbf{x}_{-i}), T)|x_i] - e < x_i$ for e small enough, which is a contradiction (notice that the strategy set of the agent with income x_i' is contained in that of the agent with income x_i). In equilibrium, agent i with income x_i is paying positive expected taxes. When he deviates and reports x_i', by definition of a taxation method, he will end up paying lower expected taxes. Notice how the assumption about the class of environments guarantees that this deviation always induces an economy. $\qquad\square$

8. Concluding remarks

For taxation problems with at least two agents, we have presented "canonical mechanisms" that implement all taxation methods (as the mechanism in normal form does) or a large class of consistent methods (as the one in extensive form). Both mechanisms appear very natural and dispense with the off-equilibrium devices that the general theory utilizes. The latter mechanism has the advantage of being suitable for environments where the information is not complete. However, no method can be implemented if the information is fully private. The implementation problem in this paper requires the tools of feasible implementation and any successful mechanism must depend on the underlying environment.

Although the assumption that agents know every income may seem restrictive, the mechanism in normal form has many advantages. In particular, it implements all taxation methods for bilateral environments and serves as a support for the mechanism in extensive form. In the latter mechanism agents may bargain over taxes with the tax collector and the interpersonal comparison of incomes is used only as a "last option." This may fit situations in which a certain legal or institutional environment surrounds multilateral negotiations in a small club, or may be representing the appeal to the central tax authority in a large decentralized economy.

It is also interesting to note that, although taxation and bankruptcy problems are treated as equivalent in the axiomatic literature, they look very different from the point of view of incentives. Formally, we denote a bankruptcy problem by $(E, \mathbf{d})$, where E is the estate to be divided and $\mathbf{d}$ is the vector of n claims on it such that $D = \sum_{i \in N} d_i \geq E$, and a bankruptcy rule $h(E, \mathbf{d})$ is a function that

assigns a feasible allocation (a division of the estate E that awards no creditor more than her claim) to each bankruptcy problem. Then, a bankruptcy problem can be translated into a taxation one as follows: the amount of taxes to be collected is $T = D - E$, that is, the money lost by the set of creditors, and the vector of taxable incomes $\mathbf{x} = \mathbf{d}$ (the claims). The taxation problem associated with the bankruptcy problem $(E, \mathbf{d})$ is $(\mathbf{d}, D - E)$. Similarly, if h is a bankruptcy rule, the corresponding taxation method would be f that assigns to each taxation problem $(\mathbf{x}, T)$ the tax allocation $f(\mathbf{x}, T) = \mathbf{x} - h(\sum_{i \in N} x_i - T, \mathbf{x})$.

From the point of view of incentives and if we maintain the same assumptions as for the taxation problem (see Sect. 3), the implementation exercise becomes trivial. That is, suppose the estate is known to the court but the true claims are not. Each claim is backed by a body of evidence. Clearly, a claimant may choose to use only part of the supporting evidence the day of the trial (over-reporting is technologically ruled out). Then, since utility is increasing in money and most bankruptcy rules are monotone in one's claim, we can solve the implementation problem in dominant strategies by asking each creditor her claim. Every creditor will report her true claim regardless of the other announcements. Notice that the dominant strategy in taxation in a similar game (to report 0 income) does not implement any taxation method. Finally, separate from the implementation problem, the reader is referred to Dagan et al. (1997) for a bargaining view of bankruptcy rules when the claims are known to the court.

References

1. Adams, C. (1993) *For Good and Evil - The Impact of Taxes on the Course of Civilization*. Madison Books, London
2. Aumann, R. J., Maschler, M. (1985) Game Theoretic Analysis of a Bankruptcy Problem from the Talmud. *Journal of Economic Theory* 36: 195–213
3. Bentham, J. (1789) *An Introduction to the Principles of Morals and Legislation*. Oxford (reprinted by Hafner, New York, 1948)
4. Bernheim, D., Peleg, B., Whinston, M. (1987) Coalition-Proof Nash Equilibria I - Concepts. *Journal of Economic Theory* 43: 1–12
5. Cassel, G. (1901) The Theory of Progressive Taxation. *Economic Journal* 11: 481–491
6. Cohen Stuart, A. J. (1889) On Progressive Taxation. In: Musgrave, R. A., Peacock, A. (eds.) *Classics in the Theory of Public Finance*. Macmillan, New York, 1958
7. Dagan, N. (1994) On the Least Sacrifice Principle in Taxation. D.P. #50, Center for Rationality and Interactive Decision Theory. Hebrew University of Jerusalem
8. Dagan, N., Serrano, R., Volij, O. (1997) A Non-Cooperative View of Consistent Bankruptcy Rules. *Games and Economic Behavior* 18: 55–72
9. Edgeworth, F. Y. (1898) The Pure Theory of Taxation. In: Musgrave, R. A., Peacock, A. (eds.) *Classics in the Theory of Public Finance*. Macmillan, New York, 1958
10. Hurwicz, L., Maskin, E., Postlewaite, A. (1995) Feasible Nash Implementation of Social Choice Rules when the Designer Does not Know Endowments or Production Sets. In: Ledyard, J. O. (ed.) *The Economics of Informational Decentralization: Complexity, Efficiency and Stability*. Kluwer Academic Publishers, Amsterdam, pp. 367–433
11. Jackson, M. (1992) Implementation in Undominated Strategies: a Look at Bounded Mechanisms. *Review of Economic Studies* 59: 757–775
12. Maskin, E. (1977) Nash Equilibrium and Welfare Optimality. Mimeo, M.I.T

13. Maskin, E. (1985) The Theory of Implementation in Nash Equilibrium: a Survey. In: Hurwicz, L., Schmeidler, D., Sonnenschein, H. (eds.) *Social Goals and Social Organization*. Cambridge University Press, Cambridge
14. Mill, J. S. (1848) *Principles of Political Economy*, (revised edition 1900). P. F. Collier and son, New York
15. Mirrlees, J. (1971) An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies* 38: 175–208
16. Mirrlees, J. (1986) The Theory of Optimum Taxation. In: Arrow, K. J., Intriligator, M. D. (eds.) *Handbook of Mathematical Economics, vol. III*. North-Holland, New York
17. Moore, J. (1992) Implementation, Contracts and Renegotiation in Environments with Complete Information. In: Laffont, J. J. (ed.) *Advances in Economic Theory, vol. I*, (VI World Congress of the Econometric Society). Cambridge University Press, Cambridge
18. O'Neill, B. (1982) A Problem of Rights Arbitration from the Talmud. *Mathematical Social Sciences* 2: 345–371
19. Piketty, T. (1993) Implementation of First-Best Allocations via Generalized Tax Schedules. *Journal of Economic Theory* 61: 23–41
20. Postlewaite, A., Wettstein, D. (1989) Feasible and Continuous Implementation. *Review of Economic Studies* 56: 603–611
21. Ricardo, D. (1817) Principles of Political Economy and Taxation. London
22. Serrano, R., Vohra, R. (1997) Non-Cooperative Implementation of the Core. *Social Choice and Welfare* 14: 513–525
23. Thomson, W. (1990) The Consistency Principle. In: Ichiishi, T., Neyman, A., Tauman, Y. (eds.) *Game Theory and Applications*. Academic Press, San Diego, CA
24. Tian, G. (1989) Implementation of the Lindahl Correspondence by a Single-Valued, Feasible and Continuous Mechanism. *Review of Economic Studies* 56: 613–621
25. Tian, G. (1993) Implementing Lindahl Allocations by a Withholding Mechanism. *Journal of Mathematical Economics* 22: 169–179
26. Tian, G. (1994) Implementation of Linear Cost Share Equilibrium Allocations. *Journal of Economic Theory* 64: 568–584
27. Young, H. P. (1987) Progressive Taxation and the Equal Sacrifice Principle. *Journal of Public Economics* 32: 203–214
28. Young, H. P. (1988) Distributive Justice in Taxation. *Journal of Economic Theory* 44: 321–335