



On the optimality of  
optimal income taxation

Felix Bierbrauer





# On the optimality of optimal income taxation

Felix Bierbrauer

May 2010

# On the optimality of optimal income taxation\*

Felix J. Bierbrauer<sup>†</sup>  
Max Planck Institute, Bonn

April 29, 2010

## Abstract

The Mirrleesian model of income taxation restricts attention to simple allocation mechanism with no strategic interdependence, i.e., the optimal labor supply of any one individual does not depend on the labor supply of others. It has been argued by Piketty (1993) that this restriction is substantial because more sophisticated mechanisms can reach first-best allocations that are out of reach with simple mechanisms. In this paper, we assess the validity of Piketty's critique in an independent private values model. As a main result, we show that the optimal sophisticated mechanism is a simple mechanism, or, equivalently, a Mirrleesian income tax system.

*Keywords:* Optimal Income Taxation, Mechanism Design

*JEL:* D82, D86, H21

---

\*I am indebted to Martin Hellwig for numerous discussions about the relationship between the theory of taxation and the theory of mechanism design. I am also grateful for conversations with Alia Gizatulina, Mike Golosov and Hans Peter Grüner.

<sup>†</sup>Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, 53113 Bonn, Germany.  
Email: bierbrauer@coll.mpg.de

# 1 Introduction

We study a model with  $n$  individuals who produce output and consume a private good. As in the Mirrleesian model of optimal income taxation, these individuals differ in their productive abilities and it is desirable that more able individuals produce more output, while marginal utilities of consumption should be equalized. Moreover, individuals are privately informed about their productive abilities. We seek to characterize the optimal mechanism which specifies, for each individual, an output requirement and a consumption level both as a function of the individual's type and the types of all other individuals. Our main result provides conditions so that (i) the optimal mechanism is a *simple mechanism* in the sense that the outcome for any one individual depends only on the characteristics of that individual, irrespectively of what the characteristics of other individuals are, and (ii) the optimal mechanism is equivalent to an optimal Mirrleesian income tax.

Our main assumption is that the productive abilities of different individuals are independent and identically distributed random variables. Also, our mechanism design approach is based on a condition of budget balance in expectation, i.e., it is required that the expected level of output per capita is not less than the expected level of consumption per capita. With a given finite number of individuals, the Mirrleesian income tax may therefore violate the public sector budget constraint, provided that budget balance holds on average. However, we show that, as the number of individuals grows without limit, the law of large numbers implies that the probability of a budget surplus or deficit converges to zero.

Thus, the main conclusion of our analysis is that, in a large economy, there is no mechanism which outperforms the Mirrleesian income tax. The result extends to a small economy if budget balance is required on average, but not if budget balance is required for all states that occur with positive probability.

This result is of interest because the existing literature is unclear with respect to the interpretation of the Mirrleesian income tax model from a mechanism design perspective. In particular, there have been interpretations based on a model with a finite number of individuals, and interpretations based on a model with a continuum of individuals, which give rise to very different conclusions.

More specifically, for a continuum economy, it has been shown by Hammond (1979) and Guesnerie (1995) that the Mirrleesian income tax problem is equivalent to a model of optimal mechanism design. For a model with a finite number of individuals, it has been shown by Dierker and Haller (1990) that the Mirrleesian problem is equivalent to a model of mechanism design in which attention is restricted to simple mechanisms. For this model, Piketty (1993) has shown that the focus on simple mechanisms involves a substantial loss of generality. First-best welfare optima, including the utilitarian one, can be reached with mechanisms that exploit the option to make the outcome for an individual  $i$  dependent on the behavior of some other individual  $j$ . These outcomes are out of reach with simple mechanisms, or, equivalently, with income tax schedules.<sup>1</sup>

---

<sup>1</sup>Hamilton and Slutsky (2007) extend Piketty's result. They show that his results remain valid if budget

Piketty’s work challenges the approach by Hammond (1979) and Guesnerie (1995). Most economists view an analysis with a continuum of individuals as useful only to the extent that it provides an approximation of the outcomes that would be obtained with a large finite economy. If the results for a finite economy (first-best is possible) are fundamentally different from the results for a continuum economy (first-best is out of reach), this casts doubt on the plausibility of a model with a continuum of individuals.

The Mirrleesian income tax problem is appealing to many researchers because – in contrast to Ramsey models of taxation which are theoretically interesting only if lump sum taxes are assumed unavailable – it limits the role of a priori assumptions on admissible policies. Now, what are the a priori assumptions in the Mirrleesian model? According to Dierker and Haller (1990) attention is restricted to simple mechanism. Piketty’s result then leads us to the conclusion that the Mirrleesian policy problem is somewhat contrived because without this restriction we could reach first-best. The Mirrleesian approach to optimal taxation is therefore subject to the same criticism as Ramsey models of taxation: The policy problem is interesting only because of ad hoc restrictions on the set of available policy instruments.

The present paper introduces an alternative model of a finite economy. With this model, the problems above do not arise. In particular, the optimal mechanism in the finite economy is shown to be a simple mechanism which is equivalent to an optimal income tax. This result does not depend on the number of individuals. The only difference between an economy with a finite number of  $n$  individuals, and the limit outcome that is obtained as  $n \rightarrow \infty$  is that, due to the law of large numbers, in the “limit economy” budget balance holds in a strict, ex post sense, whereas in the finite economy budget balance holds only in a weak, ex ante sense. This observation also implies that a model based on a large, but finite number of individuals gives approximately the same results as a model with a continuum of individuals, and therefore provides a justification for the latter approach.<sup>2</sup>

The crucial difference between our model and the one by Dierker and Haller (1990) and Piketty (1993) is as follows: Dierker and Haller (1990) and Piketty (1993) consider a finite economy in which the cross-section distribution of productive abilities is predetermined and commonly known. For instance, in an economy with two individuals, one individual is known to be high-skilled and one individual is known to be low-skilled. It is, however, unknown whether person 1 or person 2 is the high-skilled person. In particular, this implies that if the mechanism designer has learned person 1’s type, she automatically knows person 2’s type. More generally, any one individual is *informationally small* in the sense that access to the privately held information of  $n - 1$  individuals reveals the last individual’s privately held information.

In our model, by contrast, individual skill-levels are drawn independently. Knowing the type of person 1 therefore does not contain information about the type of person 2. Likewise, observing the types of  $n - 1$  individuals reveals nothing about the type of the last individual.

---

balance is also required out-of-equilibrium.

<sup>2</sup>A recent literature uses models of mechanism design with a continuum of individuals in order to characterize optimal insurance contracts or tax systems; see, for example, Golosov et al. (2003), Kocherlakota (2005), Bassetto and Phelan (2008), or Kocherlakota and Phelan (2009).

As a consequence, the cross-section distribution of types is itself a random quantity. In a model with two individuals it could well be the case that both of them turn out to be high-skilled or that both of them turn out to be low-skilled. However, we show that, as the number of individuals becomes larger, by the law of large numbers, the cross-section distribution of types converges to the probability distribution from which types are drawn. For instance, if each individual is high-skilled and low-skilled with equal probability, then, in a large finite economy, the population share of high-skilled individuals equals  $\frac{1}{2}$ , almost surely.

The fact that our information structure is different from Piketty’s does not yet imply that the restriction to simple mechanisms is without loss of generality. In fact, the optimality of simple mechanisms requires an additional assumption on the risk attitudes of individuals. The logic is as follows. With non-simple mechanisms, we can make the outcome for some individual  $i$  dependent on the types of all other individuals. From individual  $i$ ’s perspective, the types of other individuals are random. The question, then, is whether an optimal mechanism should provide full insurance of individual  $i$  against the randomness in the types of other individuals. Under an assumption of decreasing risk aversion, we show that the answer to this question is “yes”. This is the key step in our proof that simple mechanisms are optimal.

The desirability of “full insurance against the risk in other individuals’ types” can be explained as follows. When dealing with individual  $i$ , the mechanism faces an equity-efficiency trade-off. The equity concern implies that she would like  $i$ ’s consumption to be independent of  $i$ ’s type. The efficiency concern implies that  $i$  should work more if he is more productive. A necessary condition for the desirability of non-simple mechanisms is that this trade-off becomes, from the mechanism designer’s perspective, more favorable – in the sense that additional consumption of low-skilled types could be obtained at a lower welfare cost – if individual  $i$  is exposed to randomness. The trade-off would become more favorable if such randomness created slack in the incentive constraints of the very productive types of individual  $i$  so as to create room for additional consumption of the less productive types of individual  $i$ . However, with decreasing risk aversion, such randomness introduces additional slack only in the incentive constraints of the less productive types, if anything. Given that these constraints are not binding anyway, nothing is gained by having outcomes for any one individual being dependent on the outcomes of other individuals.

Our proof of these statements builds on a proof of Hellwig (2007b) who studies the desirability of randomized income taxation, or in the parlance of this paper, of stochastic simple mechanism. For the proof of our main result, we show that the arguments which imply that stochastic simple mechanisms cannot outperform deterministic simple mechanisms also imply that non-simple mechanism cannot outperform simple mechanisms.

The remainder is organized as follows. The next section describes the environment. Section 3 defines the Mirrleesian income tax problem. In Section 4, we review Piketty’s critique of the Mirrleesian model. Our main result is in Section 5. The last section contains concluding remarks. All proofs are in the Appendix.

## 2 The environment

There is a finite set of individuals, denoted by  $I = \{1, \dots, n\}$ . Individual  $i$  has a utility function

$$U_i = U(c_i, y_i, w_i) ,$$

where  $c_i$  is  $i$ 's consumption of a private good, or after-tax income,  $y_i$  is the contribution to the economy's output, or pre-tax-income, and  $w_i$  is a productivity parameter, which is private information of individual  $i$ . We assume throughout that  $U$  is twice continuously differentiable, and that the partial derivatives satisfy  $U_c > 0$ , and  $U_y < 0$ . We also assume that  $U$  is strictly quasiconcave in  $c_i$  and  $y_i$ .

For each individual  $i$ , the productivity parameter  $w_i$  belongs to a finite ordered set of possible productivity parameters  $W = \{w^1, w^2, \dots, w^m\}$ . A higher productivity parameter is associated with a lower disutility from productive effort; i.e., we assume that the single crossing condition is satisfied: For every point  $(c, y) \in \mathbb{R}_+^2$ , and every pair  $(w, w') \in W^2$  with  $w' > w$ ,

$$-\frac{U_y(c, y, w)}{U_c(c, y, w)} > -\frac{U_y(c, y, w')}{U_c(c, y, w')} .$$

Finally, we assume that, under a first-best allocation, each individual's productive effort is positive. More formally, given  $k \in \{1, \dots, m\}$ , let

$$(c^{kf}(v), y^{kf}(v)) := \operatorname{argmin}_{(c,y) \in \mathbb{R}_+^2} c - y \quad \text{s.t.} \quad U(c, y, w^k) = v .$$

Then,  $y^{kf}(v) > 0$ .

## 3 The Mirrleesian model of optimal income taxation

We introduce the Mirrleesian model of optimal income taxation as a benchmark which facilitates the interpretation of the results from the mechanism design analysis below. The Mirrleesian model is as follows: individuals choose how much to work and thereby face an income tax schedule  $T : \mathbb{R}_+ \rightarrow \mathbb{R}$  that relates their pre-tax-income to their after-tax-income; i.e., they solve the following utility maximization problem,

$$\max_{c,y} U(c, y, w_i) \quad \text{s.t.} \quad c \leq y - T(y) .$$

We denote by  $(C^*(w^i | T), Y^*(w^i | T))$  the pair which solves this utility maximization problem.

An income tax  $T$  is said to be feasible if aggregate tax revenues are non-negative,

$$\sum_{k=1}^m f^k T(Y^*(w^k | T)) \geq r ,$$

where  $f = (f^1, \dots, f^m)$  is a probability distribution with support  $W$ , and  $r$  is an exogenous revenue requirement.

At this stage, two different interpretations of the public sector budget constraint, and of the probability distribution  $f$  are possible. One interpretation is that  $f^k$  equals the population share of individuals with skill level  $w^k$ . An alternative interpretation is that  $f^k$  is the probability that any one individual has a skill level of  $w^k$ . The latter interpretation gives rise to a weaker

interpretation of the budget balance condition, namely as the requirement that expected tax revenues must be sufficient to finance the resource requirement. We will argue later that, with a large number of individuals, both interpretations are essentially equivalent.

The Mirrleesian income tax problem can now be stated as follows: choose  $(C^k, Y^k)_{k=1}^m$  in order to maximize a utilitarian social welfare function

$$S = \sum_{k=1}^m f^k U(C^k, Y^k, w^k)$$

subject to the constraints that there is a function  $T$  so that, for every  $k$ ,

$$(C^k, Y^k) = (C^*(w^k | T), Y^*(w^k | T)), \quad (1)$$

and

$$\sum_{k=1}^m f^k (Y^k - C^k) \geq r. \quad (2)$$

The taxation principle<sup>3</sup> implies that the constraints in (1) can be written in a way that no longer involves an explicit reference to a tax schedule  $T$ . According to this result,  $(C^k, Y^k)_{k=1}^m$  satisfies (1), for some income tax schedule  $T$ , if and only if it satisfies the following *Mirrleesian incentive compatibility constraints*: For every pair  $k, l$ ,

$$U(C^k, Y^k, w^k) \geq U(C^l, Y^l, w^k). \quad (3)$$

**Assumption 1** (*Non-decreasing marginal costs of public funds*)

Let  $S^*(r)$  be the level of welfare that is induced by a solution to the Mirrleesian income tax problem. We assume that, for all  $r$ ,  $S^{*''}(r) \leq 0$ .

A marginal increase of the revenue requirement  $r$  in the government budget constraint yields a welfare loss with absolute value equal to  $-S^{*'}(r)$ . Assumption 1, which will be important for the proof of our main result, says that this welfare loss is a non-decreasing function of  $r$ .<sup>4</sup>

## 4 Piketty's critique of the Mirrleesian model

In the following, we will briefly review the critique of the Mirrleesian income tax model that is due to Piketty (1993). Piketty's work is based on the interpretation of the probability distribution  $f$  as a commonly known cross-section distribution of productive abilities.

To illustrate his approach we focus on a simple setup with two individuals,  $I = \{1, 2\}$ , two possible skill levels  $W = \{w^1, w^2\}$  and a given cross-section distribution of productivity parameters  $f = (\frac{1}{2}, \frac{1}{2})$ . Note that the information structure is such that, from an outsider's perspective, one individual is known to be high-skilled and one individual is known to be low-skilled. However, the outsider does not know whether individual 1 or individual 2 is the high-skilled individual.

<sup>3</sup>See Hammond (1979) or Guesnerie (1995) for a proof.

<sup>4</sup>Assumption 1 could be traced back to assumptions about the primitives of the model, in particular of the utility function  $U$ ; see Bierbrauer and Boyer (2010) for an example. Here, this would lead us astray.

For simplicity, we assume that individuals have a separable utility function  $U(c, y, w) = u(c) - v\left(\frac{y}{w}\right)$ , where  $u$  is increasing and concave function and  $v$  is increasing and convex. Under these assumptions, the first-best utilitarian allocation consists of an consumption-output combination for the low-skilled individual,  $A^{1*} = (C^{1*}, Y^{1*})$ , and one for the high-skilled individual  $A^{2*} := (C^{2*}, Y^{2*})$  which satisfies the following properties: (i) marginal utilities of consumption are equalized, implying that  $C^{1*} = C^{2*} =: C^*$ , and (ii) for each individual, the marginal utility of consumption equals the marginal disutility of effort so that  $u'(C^*) = \frac{1}{w^1} v'\left(\frac{Y^{1*}}{w^1}\right) = \frac{1}{w^2} v'\left(\frac{Y^{2*}}{w^2}\right)$ . In particular, this implies that the high-skilled individual generates more output than the low-skilled individual:  $Y^{2*} > Y^{1*}$ .

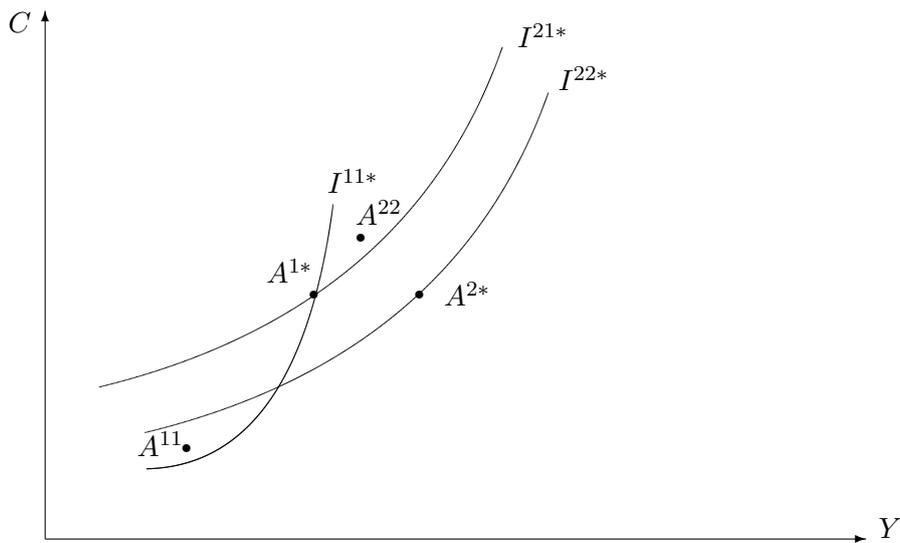
Piketty's critique of the Mirrleesian optimal income tax problem is the following: this first-best allocation is not *Mirrleesian incentive compatible*.<sup>5</sup> However, it can be reached with a more general mechanism design approach. To demonstrate this, consider the following mechanism: Each individual sends a message from the set  $W = \{w^1, w^2\}$ . Let  $(\hat{w}_1, \hat{w}_2)$  be a typical pair of messages by individuals 1 and 2, and let  $A_i(\hat{w}_1, \hat{w}_2)$  be the consumption-effort pair that individual  $i$  gets as a function of the message profile.

Suppose these consumption-effort pairs satisfy the following properties: (i) If one individual declares to be high-skilled and one individual declares to be low-skilled, the former gets the bundle  $A^{1*}$ , and the latter gets the bundle  $A^{2*}$ , (ii) if both individuals declare to be low-skilled both individuals get the bundle  $A^{11}$ , (iii) if both individuals declare to be high-skilled both individuals get the bundle  $A^{22}$ . It is easy to verify that, if  $A^{11}$  and  $A^{22}$  are chosen as in Figure 1, then truth-telling is a dominant strategy, which implies that in equilibrium the first-best allocation is reached.

More generally, Piketty shows that for an economy with  $n$  individuals and a given cross-section distribution of types, one can construct a game so that the first-best utilitarian allocation is the unique equilibrium that survives the iterated elimination of dominated strategies. His approach uses the possibility to make outcomes for one individual dependent on the behavior of the other individual. To see this, note, e.g., that  $A_1(w^1, w^2) = A^{1*} \neq A_1(w^1, w^1) = A^{11}$ . Now, if we impose the restriction that such an interdependence must not arise, then there is a bundle  $A^1 = (C^1, Y^1)$  that an individual gets whenever it communicates a low-skill level, irrespectively of what the other individual communicates, and a bundle  $A^2 = (C^2, Y^2)$  that an individual gets whenever it communicates a high-skill level. Following Dierker and Haller (1990), we refer to mechanisms with this property in the following as *simple*. It is straightforward to verify that, for the game induced by such a simple mechanism, truth-telling is a dominant strategy equilibrium if and only if  $(C^k, Y^k)_{k=1}^2$  satisfies the Mirrleesian incentive compatibility constraints in (3).

---

<sup>5</sup>To see this, note that a high-skilled individual prefers the bundle intended for the low-skilled individual,  $(C^*, Y^{1*})$ , over  $(C^*, Y^{2*})$ . The reason is that the second bundle involves a larger workload. Hence, for  $(C^1, Y^1) = (C^*, Y^{1*})$  and  $(C^2, Y^2) = (C^*, Y^{2*})$ , the inequalities in (3) are violated.



**Figure 1.** The indifference curve of a low-skilled/ high-skilled individual through the bundle  $A^{1*}$  is denoted  $I^{11*} / I^{21*}$ . The indifference curve of a high-skilled individual through  $A^{2*}$  is denoted by  $I_2^{22*}$ .

Piketty's result suggests that the Mirrleesian model of optimal income taxation lacks a theoretical foundation. Given that superior mechanisms exist, it is ultimately unclear why, for the purposes of a normative analysis, attention should be restricted to simple mechanisms.

Our plan in the following is to look at this problem from a different angle. In particular, we introduce an alternative model of a finite economy. The essential difference to Piketty's model is the following: in Piketty's analysis, each individual is informationally small. Given that the cross-section distribution  $f$  is assumed to be known, if the mechanism designer has already gained access to the privately held information of  $n - 1$  individuals about their productive abilities, he can infer the last individual's productivity level. By contrast, we will assume that the productivity level of any one individual  $i$  can not be inferred from information about the productivity levels of all other individuals.

With this model we can show that, with an arbitrary, finite number of  $n$  individuals, the optimal mechanism satisfying a feasibility condition and a Bayes-Nash incentive compatibility condition is equivalent to an optimal income tax in the Mirrlees-model. Our constraints are weaker than those imposed by Piketty. We use a weaker notion of incentive compatibility, and a requirement of budget balance in expectation rather than an ex post budget balance condition. This makes the observation that the optimal mechanism is equivalent to an optimal income tax even more striking.

## 5 A finite economy with independent private values

We study the environment in Section 2 under additional assumptions. First, we assume that the individuals' productivity levels are realizations of independent and identically distributed random variables with distribution  $f = (f^1, \dots, f^k)$ ; i.e.,  $f^k$  is now interpreted as the probability

of the event that  $w_i = w^k$ , for any one individual  $i$ . Second, we make the following assumptions about preferences:

**Assumption 2** (*Desirability of Redistribution*)

- a) The utility function  $U$  is concave in  $y$  and satisfies  $U_y(c, y, w^l) > U_y(c, y, w^k)$  for  $l > k$ .
- b) Let  $l > k$ . Let  $(c^l, y^l) \in \mathbb{R}_+^2$  and  $(c^k, y^k) \in \mathbb{R}_+^2$  be such that

$$c^l > c^k, \quad U_c(c^l, y^l, w^l) \geq -U_y(c^l, y^l, w^l) \quad \text{and} \quad U_c(c^k, y^k, w^k) = -U_y(c^k, y^k, w^k).$$

Then,  $-U_y(c^l, y^l, w^l) < -U_y(c^k, y^k, w^k)$ .

**Assumption 3** (*Decreasing Risk Aversion*)

Let  $\tilde{c}$ , and  $\tilde{y}$  be a pair of nondegenerate nonnegative-valued random variables. Let  $l > k$ . For any  $\pi \in \mathbb{R}$ , let  $E[U(\tilde{c}, \tilde{y}, w^k)] = U(E[\tilde{c}] - \pi, E[\tilde{y}], w^k)$ . Then,  $\pi > 0$  and  $E[U(\tilde{c}, \tilde{y}, w^l)] \geq U(E[\tilde{c}] - \pi, E[\tilde{y}], w^l)$ .

Assumption 2 describes situations in which it is desirable to change output requirements so that high-skilled individuals work more and low-skilled individuals work less.<sup>6</sup> Part a) states that if a high-skilled and a low-skilled individual are pooled so that they both consume the same and provide the same output, then utilitarian welfare goes up if output requirements are rearranged. Part b) extends this to first-best allocations with the property that high-skilled individuals consume more than low-skilled individuals, and to second-best allocation where the high-skilled individual's marginal disutility of output provision is even lower than with an undistorted first-best allocation.

Assumption 3 formalizes a concept of decreasing risk aversion, with risk premia measured in units of the consumption good.<sup>7</sup> The assumption says that, if an individual of some given skill level is indifferent between a lottery over consumption-output bundles and a deterministic consumption-output-bundle – in which the output component is equal to the expected level of output under the lottery –, then an individual with a high-skill level will not prefer the deterministic bundle over the lottery.

Hellwig (2007b) shows that these two assumptions have the following implications. First, for the Mirrleesian income tax problem defined in Section 3, one may without loss of generality limit attention to a subset of the Mirrleesian incentive compatibility constraints, namely to the Mirrleesian *downward* incentive compatibility constraints: for every pair  $k$  and every  $h < k$ ,

$$u(C^k, Y^k, w^k) \geq u(C^h, Y^h, w^k). \tag{4}$$

A solution of the relaxed Mirrleesian problem which takes only these constraints into account satisfies the neglected upward incentive compatibility constraints automatically. Second, the

---

<sup>6</sup>For a more extensive discussion of this assumption and its relation to alternative assumptions that have been made in the literature on optimal income taxation, see Hellwig (2007a).

<sup>7</sup>See Hellwig (2007b) for more details.

option to assign stochastic rather than deterministic consumption-output-bundles to individuals, does not yield a higher level of utilitarian welfare and will hence not be used.

## 5.1 The mechanism design problem

Our main result will be that the optimal Mirrleesian income tax solves a mechanism design problem with no a priori restrictions on the set of admissible mechanisms. Before we can state this result, we first need to define this mechanism design problem.

We appeal to the revelation principle and limit attention to direct mechanisms that possess truth-telling equilibria. We use Bayes-Nash equilibrium as the solution concept. Let  $w = (w_1, \dots, w_n)$  be a profile of productivity levels. A direct mechanism specifies a pair of outcome functions  $c_i : w \mapsto c_i(w)$  and  $y_i : w \mapsto y_i(w)$ , for each individual  $i \in I$ . Truth-telling is a Bayes-Nash equilibrium provided that the following *Bayes-Nash incentive compatibility constraints* are satisfied: For each  $i$ , for each  $k$  and for each  $l$ ,

$$E_{w_{-i}}[U(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k), w^k)] \geq E_{w_{-i}}[U(c_i(w_{-i}, w^l), y_i(w_{-i}, w^l), w^k)]. \quad (5)$$

A mechanism is said to be feasible if the expected budget surplus is non-negative,

$$E_w \left[ \sum_{i=1}^n (y_i(w) - c_i(w)) \right] \geq 0. \quad (6)$$

This budget condition is weaker than one requiring budget balance in an ex post sense, so that, for every  $w$ ,  $\sum_{i=1}^n (y_i(w) - c_i(w)) \geq 0$ . However, we will show below (see Section 5.4), that, under an optimal mechanism, the probability of the event  $\sum_{i=1}^n (y_i(w) - c_i(w)) \neq 0$  converges to zero, as  $n \rightarrow \infty$ .

The mechanism design problem is to choose the functions  $(c_i)_{i=1}^n$  and  $(y_i)_{i=1}^n$  in order to maximize expected utilitarian welfare,  $ES := E_w [\sum_{i=1}^n U(c_i(w), y_i(w), w_i)]$ , subject to the constraints in (5) and (6).

## 5.2 The main result

### Proposition 1

- i) The optimal mechanism is a simple mechanism. Let  $(c_i^*)_{i=1}^n$  and  $(y_i^*)_{i=1}^n$  be the solution to the mechanism design problem in Section 5.1. Then, for every  $k \in \{1, \dots, m\}$ , there exists  $(C^{k*}, Y^{k*}) \in \mathbb{R}_+^2$  so that, for every  $i$ ,  $w_i = w^k$  implies that*

$$(c^{i*}(w_{-i}, w^k), y^{i*}(w_{-i}, w^k)) = (C^{k*}, Y^{k*}),$$

*for every  $w_{-i} \in W^{n-1}$ .*

- ii) The optimal mechanism is an optimal income tax. The collection  $(C^{k*}, Y^{k*})_{k=1}^m$  solves the Mirrleesian income tax problem in Section 3.*

The proof of the main result builds on the proof in Hellwig (2007b) that random income tax schedules cannot outperform deterministic ones. More specifically, we first focus on a relaxed mechanism design problem where, for each individual  $i$ , only *downward* Bayes-Nash incentive compatibility constraints are imposed. Given that these are the only incentive constraints that are taken into account, we show that non-simple mechanisms cannot outperform simple mechanisms.

The key step in the proof of this statement is the following. Any individual  $i$  is willing to pay a risk premium for being insured against the fluctuations in  $w_{-i}$ . Hence, if we start from a non-simple mechanism that is downward Bayes-Nash incentive compatible and replace it by a simple mechanism so that every individual's expected output level remains the same and every individuals' expected utility level remains the same, we can reduce expected consumption levels and create a surplus in the feasibility constraint (6). Moreover, because of decreasing risk aversion, type  $w^l$  of individual  $i$  considers the consumption-output-combination for type  $w^k$  under the new, simple mechanism to be less attractive than under the old, non-simple mechanism. Hence, the newly constructed simple mechanism is also downward Bayes-Nash incentive compatible.

This implies that, as long as we limit attention to downward incentive constraints, focussing on simple mechanisms involves no further loss of generality. Given that with simple mechanisms outcomes for different individuals are linked only via the budget constraint and that simple mechanisms are equivalent to income tax schedules, the relaxed mechanism design problem can now be decomposed into a number of subproblems: first, there is a separate subproblem for each individual  $i$ : choose a simple mechanism  $(C_i^k, Y_i^k)_{k=1}^n$  that satisfies the downward *Mirrleesian* incentive constraints and yields an expected tax revenue of  $\sum_{k=1}^n f^k(Y_i^k - C_i^k) = T_i$ . Second, choose the expected tax payments of different individuals  $(T_i)_{i=1}^n$  in a welfare-maximizing way, subject to the constraint that  $\sum_{i=1}^n T_i \geq 0$ .

An individual's subproblem is in fact a relaxed Mirrleesian income tax problem with a revenue requirement of  $T_i$ . It follows from the analysis of the relaxed Mirrleesian income tax problem in Hellwig (2007b) that the solution to each individual's subproblem is also upward incentive compatible. This proves that the optimal simple mechanism that is downward incentive compatible, is also upward incentive compatible, and therefore is a solution to the "full" mechanism design problem.

The last step in the proof is to verify that the optimal simple mechanism is in fact symmetric or anonymous. A utilitarian mechanism designer chooses the revenue requirements of different individuals so that the marginal impact on expected utility levels is equalized. Assumption 1 implies that this is achieved only if each individual faces the same revenue requirement.

### 5.3 Robustness and Dominant Strategies

The optimal mechanism satisfies the Bayes-Nash incentive compatibility constraints. However, the fact that the optimal mechanism is a simple mechanism implies that it also satisfies the more demanding *ex post incentive compatibility constraints*: For each  $i$ , for each  $k$ , for each  $l$ ,

and for each  $w_{-i}$ ,

$$U(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k), w^k) \geq U(c_i(w_{-i}, w^l), y_i(w_{-i}, w^l), w^k) . \quad (7)$$

These constraints can be interpreted as follows: suppose that individual  $i$  has learned what the types of the other individuals are. Then, whatever these types are, individual  $i$ 's best response still is to communicate the own type truthfully to the mechanism designer. Put differently, from an ex post perspective where the individuals' private information has become public, no type of individual  $i$  regrets to have revealed himself to the mechanism designer.

A simple mechanism trivially satisfies these constraints, because the outcome for individual  $i$  is the same irrespectively of what the types of other individuals are. In fact, this implies that truth-telling is a dominant strategy. Consequently, the results of our analysis would remain unaffected if, for the mechanism design problem in Section 5.1, we replaced the Bayes-Nash incentive constraints by the more demanding incentive constraints in (7).

It has been shown by Ledyard (1978) and Bergemann and Morris (2005) that ex post incentive compatibility is a necessary and sufficient condition for the robust implementability of an allocation rule, or social choice function. Robustness here refers to the specification of the individuals' probabilistic beliefs about the environment; that is, an outcome is robustly implementable if we can implement it whatever the probabilistic beliefs of individuals look like. We chose a particular specification of beliefs, namely the one which is implied by a common prior according to which the types of different individuals are independent and identically distributed random variables. The observation that the optimal mechanism is ex post incentive compatible implies that these assumptions can be substantially weakened. In fact, every possible specification of the individuals' beliefs would give rise to the same result, provided that the mechanism designer's beliefs remain unaltered.

## 5.4 A large economy

Our main result in Proposition 1 shows that, for an economy with an arbitrary, but finite number of individuals, the optimal mechanism is equivalent to an optimal Mirrleesian income tax. This result is based on the interpretation of  $f$  as a probability distribution from which the individuals' types are drawn. This implies, in particular, that the budget constraint in the Mirrleesian model is interpreted as a constraint which requires that the budget is balanced in expectation, but not necessarily for each possible type profile. The following Proposition uses the law of large numbers in order to show that, in a large economy, we can interpret  $f$  also as the the empirical cross-section distribution of types in the economy as a whole.

**Proposition 2** *Let  $P^n(x)$  be the probability of an event  $x$  in an economy with  $n$  individuals. Let  $\mu_n^k(w) = \frac{\#\{i|w_i=w^k\}}{n}$ . For every  $\varepsilon > 0$ , and every  $\delta > 0$ , there exists  $N$  so that  $n > N$  implies*

$$P^n(|\mu_n^1(w) - f^1| < \varepsilon, |\mu_n^2(w) - f^2| < \varepsilon, \dots, |\mu_n^m(w) - f^m| < \varepsilon) \leq \delta .$$

An immediate implication of this Proposition is that, in an arbitrarily large, finite economy, ex post budget balance holds almost surely. This is stated formally in the following Corollary.

**Corollary 1** Let  $P^n(IM(\varepsilon))$  be the probability that there is an imbalance larger than  $\varepsilon$ ,

$$\begin{aligned} P^n(IM(\varepsilon)) &:= P^n(|\sum_{k=1}^m \mu_n^k(w)(C^{k*} - Y^{k*})| > \varepsilon) \\ &= P^n(|\sum_{k=1}^m \mu_n^k(w)(C^{k*} - Y^{k*}) - \sum_{k=1}^m f^k(C^{k*} - Y^{k*})| > \varepsilon). \end{aligned}$$

Proposition 2 implies that, for every  $\varepsilon > 0$ ,  $\lim_{n \rightarrow \infty} P^n(IM(\varepsilon)) = 0$ .

## 6 Concluding Remarks

In this paper, we have provided a rationale for the use of simple mechanisms in problems of social insurance, or redistributive income taxation. Simple mechanisms separate the provision of working incentives to any one individual from the provision of working incentives to other individuals. The main result of the paper shows that, under an assumption of decreasing risk aversion, the optimal mechanism is a simple mechanism. Moreover, the optimal simple mechanism is equivalent to an optimal income tax in the Mirrleesian model of income taxation.

Versions of this result are likely to carry over to other allocation problems. We illustrate this by means of two examples. First, consider a monopolistic firm who sells a private good to  $n$  consumers with private information on their willingness to pay. With a general mechanism design approach it is possible to make the quantity sold to and the price paid by a consumer  $i$  dependent on the willingness to pay of some other consumer  $j$ . A straightforward adaptation of the arguments in this paper show that if (i) the firm has a constant returns to scale technology and, (ii) agents with a low willingness to pay are not more risk-loving than agents with a high willingness to pay, then this option will not be attractive for a profit-maximizing firm. Instead, the firm will use a simple pricing mechanisms.

Second, consider an organization with  $n$  risk-averse agents/ employees and one risk-neutral principal/ employer. For each agent, the principal observes a performance measure which is a noisy signal of the agent's effort. The question now is whether the wage contract of agent  $i$  should be simple in the sense that the wage paid to  $i$  depends only on  $i$ 's performance, or whether there is a role for non-simple wage contracts that make the payment to  $i$  also dependent on  $j$ 's performance. Our analysis suggests that, if (i) the signals of different agents are stochastically independent, and (ii) agents with bad outcomes are not more risk-loving than agents with good outcomes, than the use of simple wage contracts will be optimal.

## References

- 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. and Boyer, P. (2010). The Pareto-frontier in a simple Mirrleesian model of income taxation. Mimeo, Max Planck Institute for Research on Collective Goods.
- Dierker, E. and Haller, H. (1990). Tax systems and direct mechanisms in large finite economies. *Journal of Economics*, 52:99–116.

- Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). Optimal indirect and capital taxation. *Review of Economic Studies*, 70:569–587.
- Guesnerie, R. (1995). *A Contribution to the Pure Theory of Taxation*. Cambridge University Press.
- Hamilton, J. and Slutsky, S. (2007). Optimal nonlinear income taxation with a finite population. *Journal of Economic Theory*, 132:548–556.
- Hammond, P. (1979). Straightforward individual incentive compatibility in large economies. *Review of Economic Studies*, 46:263–282.
- Hellwig, M. (2007a). A contribution to the theory of optimal utilitarian income taxation. *Journal of Public Economics*, 91:1449–1477.
- Hellwig, M. (2007b). The undesirability of randomized income taxation under decreasing risk aversion. *Journal of Public Economics*, 91:791–816.
- Kocherlakota, N. (2005). Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation. *Econometrica*, 73:1587–1621.
- Kocherlakota, N. and Phelan, C. (2009). On the robustness of laissez-faire. *Journal of Economic Theory*, forthcoming.
- Ledyard, J. (1978). Incentive compatibility and incomplete information. *Journal of Economic Theory*, 18:171–189.
- Piketty, T. (1993). Implementation of first-best allocations via generalized tax schedules. *Journal of Economic Theory*, 61:23–41.

## A Appendix

### A.1 Proof of Proposition 1

**Step 1: The relaxed mechanism design problem.** Instead of looking at directly at the mechanism design problem in Section 5.1, we first study a *relaxed mechanism design problem* that takes only local downward incentive compatibility constraints into account. More precisely, we consider the problem of choosing  $(c_i)_{i=1}^n$  and  $(y_i)_{i=1}^n$  in order to maximize  $ES$  subject to the feasibility constraint in (6) and the following subset of all *Bayes-Nash incentive compatibility constraints*: For all  $i$ , all  $k > 1$ , and all  $h < k$ ,

$$E_{w_{-i}}[U(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k), w^k)] \geq E_{w_{-i}}[U(c_i(w_{-i}, w^h), y_i(w_{-i}, w^h), w^k)] .$$

**Lemma 1** *Let  $(c_i, y_i)_{i=1}^n$  be a feasible and locally downward incentive compatible mechanism. Suppose there is  $i$ ,  $w^k$ ,  $w_{-i}$  and  $w'_{-i}$  so that*

$$(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k)) \neq (c_i(w'_{-i}, w^k), y_i(w'_{-i}, w^k)) .$$

Then, there exists a collection of simple mechanisms  $(C, Y) := (C_i, Y_i)_{i=1}^n$ , where  $(C_i, Y_i) = (C_i^k, Y_i^k)_{k=1}^m$ , that satisfies the following properties:

i) The simple mechanisms are payoff equivalent: For all  $i$  and  $k$ ,

$$U(C_i^k, Y_i^k, w^k) = E_{w_{-i}}[U(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k), w^k)].$$

ii) The simple mechanisms are locally downward incentive compatible: For all  $i$ ,  $k > 1$ , and  $h < k$ ,

$$U(C_i^k, Y_i^k, w^k) \geq U(C_i^h, Y_i^h, w^k).$$

iii) The simple mechanisms require less resources,

$$\sum_{i=1}^n \sum_{k=1}^m f^k(Y_i^k - C_i^k) < E \left[ \sum_{i=1}^n y_i(w) - c_i(w) \right].$$

**Proof** As a first step, we only adjust the mechanism for individual  $i$ , leaving  $(c_j, y_j)$  unaffected for  $j \neq i$ . Suppose there is a type  $w^k$  of individual  $i$  and a pair  $w_{-i}$  and  $w'_{-i}$  so that

$$(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k)) \neq (c_i(w'_{-i}, w^k), y_i(w'_{-i}, w^k)).$$

Define  $Y_i^k$  by the equation  $Y_i^k := E_{w_{-i}}[y_i(w_{-i}, w^k)]$ , and define  $C_i^k$  by the equation

$$U(C_i^k, Y_i^k, w^k) = E_{w_{-i}}[U(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k), w^k)]. \quad (8)$$

Construct a new mechanism  $(\bar{c}_i, \bar{y}_i)$  for individual  $i$  so that,  $\bar{c}_i(w) = C_i^k$  and  $\bar{y}_i(w) = Y_i^k$ , whenever  $w_i = w^k$ . Whenever  $w_i \neq w^k$ , let  $\bar{c}_i(w) = c_i(w)$  and  $\bar{y}_i(w) = y_i(w)$ . By construction, the new mechanism satisfies payoff equivalence. Because of decreasing risk aversion it also satisfies local downward incentive compatibility: Equation (8) in conjunction with the assumption of decreasing risk aversion implies that,

$$U(C_i^k, Y_i^k, w^l) \leq E_{w_{-i}}[u(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k), w^l)], \quad (9)$$

for all  $l > k$ . Moreover, by local downward incentive compatibility of  $(c_i, y_i)$

$$E_{w_{-i}}[U(c_i(w_{-i}, w^k), y_i(w_{-i}, w^k), w^l)] \leq E_{w_{-i}}[U(c_i(w_{-i}, w^l), y_i(w_{-i}, w^l), w^l)]. \quad (10)$$

Consequently, the new allocation is also locally downward incentive compatible. Finally, decreasing risk aversion implies that  $C_i^k < E_{w_{-i}}[c_i(w_{-i}, w^k)]$ .

We can repeat this argument for every possible type of individual  $i$ . This yields a simple mechanism  $(C_i, Y_i) = (C_i^k, Y_i^k)_{k=1}^m$  satisfying payoff equivalence, downward incentive compatibility,  $Y_i^k = E_{w_{-i}}[y_i(w_{-i}, w^k)]$ , for all  $k$  and  $C_i^k \leq E_{w_{-i}}[c_i(w_{-i}, w^k)]$ , for all  $k$ , with a strict inequality for some  $k$ . Hence,  $\sum_{k=1}^m f^k[Y_i^k - C_i^k] < E_w[y_i(w) - c_i(w)]$ .

Finally, to establish the Lemma, we repeat these arguments for all individuals. ■

The proof of Lemma 1 is an adaptation of the proof of Lemma 5.2 in Hellwig (2007b) to the given setup. It implies that for the analysis of the relaxed mechanism design problem we may focus without loss of generality on simple mechanism. We therefore define the following *relaxed simple mechanism design problem*: choose a collection of simple mechanisms  $(C_i, Y_i)_{i=1}^n$ , where  $(C_i, Y_i) = (C_i^k, Y_i^k)_{k=1}^m$ , in order to maximize

$$ES = \sum_{i=1}^n \sum_{k=1}^m f^k U(C_i^k, Y_i^k, w^k)$$

subject to the downward incentive compatibility constraints, for all  $i, k > 1$  and  $h < k$ ,

$$U(C_i^k, Y_i^k, w^k) \geq U(C_i^h, Y_i^h, w^k), \quad (11)$$

and the feasibility constraint,

$$\sum_{i=1}^n \sum_{k=1}^m f^k (Y_i^k - C_i^k) \geq 0. \quad (12)$$

The solution to this problem will henceforth be denoted by  $(C_i^s, Y_i^s)_{i=1}^n$ .

An interdependence of the simple mechanisms for two individuals  $i$  and  $j$  arises only via the feasibility constraint. For each individual  $i$ , it must therefore be true that the simple mechanism  $(C_i^s, Y_i^s) = (C_i^{ks}, Y_i^{ks})_{k=1}^m$  solves the following problem: Maximize

$$\sum_{k=1}^m f^k U(C_i^k, Y_i^k, w^k)$$

subject to the downward incentive constraints in (11) and the constraint that

$$\sum_{k=1}^m f^k (Y_i^k - C_i^k) \geq T_i, \quad (13)$$

where  $T_i := \sum_{k=1}^m f^k (Y_i^{ks} - C_i^{ks})$  is the expected payment of individual  $i$ . We refer to this problem in the following as *individual  $i$ 's relaxed simple mechanism design problem*. We denote the expected utility level of individual  $i$  that results from a solution to this problem by  $V_i(T_i)$ .

Individual  $i$ 's relaxed simple mechanism design problem is equivalent to the “*deterministic relaxed income tax problem*” studied in Hellwig (2007b). It is shown in Hellwig (2007b) that, at a solution to  $i$ 's problem, the constraint (13) holds as an equality. Any slack in this budget constraint could be used to increase individual  $i$ 's expected utility in an incentive compatible way. This observation also implies that the budget constraint (12) for the *relaxed simple mechanism design problem* holds as an equality. Otherwise, it was possible to reduce the expected payment of some individual  $i$ , and use the additional resources to make this individual better off.

These observations imply that we can characterize the solution to the relaxed simple mechanism design problem via the following two stage procedure: *first*, we solve for each individual  $i$  the corresponding simple relaxed mechanism design problem, treating the expected transfer payments  $(T_i)_{i=1}^n$  as given parameters. *Second*, we solve the *transfer problem*, i.e., we choose  $(T_i)_{i=1}^n$  in order to maximize  $\sum_{i=1}^n V_i(T_i)$  subject to  $\sum_{i=1}^n T_i = 0$ .

**Step 2: Show that the solution to the relaxed mechanism design problem and the solution of the mechanism design problem coincide.** Given that Assumptions 2 and 3 are satisfied, it is shown in Hellwig (2007b) that the solution to the *deterministic relaxed income tax problem* is incentive compatible, and is hence a solution to the Mirrleesian income tax problem. For our setting, this implies that a solution to *individual  $i$ 's relaxed simple mechanism design problem* is not only downward incentive compatible, but also upward incentive compatible, i.e., it satisfies for all  $k \geq 1$ , and all  $l$  with  $m \geq l > k$ ,

$$U(C_i^k, Y_i^k, w^k) \geq U(C_i^l, Y_i^l, w^k). \quad (14)$$

This implies, in particular, that the solution to the relaxed mechanism design problem is upward incentive compatible. It also proves that the solution to the mechanism design problem in Section 5.1 is a simple mechanism.

**Step 3: Show that the solution to the relaxed mechanism design problem and the income tax problem coincide.** Step 2 shows that each individual's relaxed simple mechanism design problem solves a Mirrleesian income tax problem. Hence, we have that  $V_i(T_i) = S^*(T_i)$ . We can therefore rewrite the transfer problem as follows: choose  $(T_i)_{i=1}^n$  in order to maximize  $\sum_{i=1}^n S^*(T_i)$  subject to  $\sum_{i=1}^n T_i = 0$ . Under Assumption 1, the solution to this transfer problem involves  $S^{*'}(T_i) = S^{*'}(T_j)$ , and hence  $T_i = T_j$ , for any pair of individuals  $i$  and  $j$ . The solution to the transfer problem therefore is such that, for all  $i$ ,  $T_i = 0$ .

Hence, each individual's relaxed simple mechanism design problem solves a Mirrleesian income tax problem with a revenue requirement of 0. This implies in particular, that the optimally chosen simple mechanisms for different individuals are identical, and coincide with the solution to the Mirrleesian income tax problem in Section 3.

## A.2 Proof of Proposition 2

We have that

$$\begin{aligned} P^n(|\mu_n^1(w) - f^1| > \varepsilon, |\mu_n^2(w) - f^2| > \varepsilon, \dots) &\leq \sum_{k=1}^m P^n(|\mu_n^k(w) - f^k| > \varepsilon) \\ &\leq m \max_k P^n(|\mu_n^k(w) - f^k| > \varepsilon). \end{aligned} \quad (15)$$

Observe that  $\mu^k(w)$  is the average of a sequence of  $n$  identically distributed and independent random variables:

$$\mu^k(w) = \frac{\sum_{i=1}^n \mathbf{1}(w_i = w^k)}{n},$$

where  $\mathbf{1}(w_i = w^k)$  is an indicator variable that takes the value 1 if  $w_i = w^k$ , and takes the value 0 otherwise. Also note that  $E[\mu^k(w)] = f^k$ . By the strong law of large numbers, for every  $k$ , and every  $\tilde{\delta} > 0$ , there exists  $\tilde{N}$  so that  $n > \tilde{N}$  implies that

$$P^n(|\mu_n^k(w) - f^k| > \varepsilon) \leq \tilde{\delta}. \quad (16)$$

Now, if let  $\tilde{\delta} = \frac{\delta}{m}$  and use (16) to substitute for  $\max_k P^n(|\mu_n^k(w) - f^k| > \varepsilon)$  in (15), we obtain

$$P^n(|\mu_n^1(w) - f^1| > \varepsilon, |\mu_n^2(w) - f^2| > \varepsilon, \dots, |\mu_n^m(w) - f^m| > \varepsilon) \leq \delta.$$

■