Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2005/25

Optimal Income Taxation and Public Good Provision in a Two-Class Economy

Felix Bierbrauer

MAX PLANCK SOCIETY

Preprints of the
Max Planck Institute
for Research on Collective Goods Bonn 2005/25

Optimal Income Taxation and Public Good Provision in a Two-Class Economy

Felix Bierbrauer

November 2005

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
http://www.mpp-rdg.mpg.de

Optimal Income Taxation and Public Good Provision in a Two-Class Economy

Felix Bierbrauer*

Max Planck Institute for Research on Collective Goods, Bonn
and CDSEM, University of Mannheim

November 9, 2005

Abstract

This paper combines the problem of optimal income taxation with the free-rider problem in public good provision. There are two groups of individuals with private information on their earning ability and their valuation of a public good. Adjustments of the transfer system are needed to discourage the more productive from exaggerating the desirability of public good provision. Similarly, the less productive need to be prevented from understating their valuation. Relative to an optimal income tax, which focuses solely on earning ability, income transfers are increased whenever a public good is installed and are decreased otherwise.

Keywords: Income Taxation, Public Good Provision, Revelation of Preferences, Two-dimensional Heterogeneity.

JEL: D71, D82, H21, H41

1 Introduction

This paper combines the problem of optimal income taxation with the free-rider problem in public good provision. An optimal income tax is based on the utilitarian desire to redistribute resources in favor of the less able. An optimal solution of the free-rider problem has the property that a public good is installed if and only if the aggregate valuation in the economy is

*I am very grateful for numerous discussions with and encouragement from Martin Hellwig. I also benefited a lot from comments of Christoph Engel, Thomas Gaube, Hendrik Hakenes, Marco Sahm and Ingolf Schwarz. I thank participants of the ENTER Jamboree 2004 in Barcelona, the conference in tribute to J.-J. Laffont in Toulouse, 2005, as well as seminar participants at the universities of Mannheim, Cologne and Munich.

sufficiently high. The present paper studies the interaction between these problems. It arises because expenditures on public goods and on income transfers are linked through a public sector budget constraint. That is, they compete for the same funds. Consequently, an individual's view on the desirability of public good provision will depend on the way he is treated by the transfer system.

To illustrate this, consider a "welfare state", which allocates a lot of resources to a transfer system. Obviously, the beneficiaries of this system are individuals with a rather low level of income. Suppose that the magnitude of the transfer system depends on the level of public good provision. That is, whenever tax revenues are used for a public good, there are less funds left for transfers. Now consider asking a person with a high income about her views on using public money for a certain project, say a highway or an opera building. As she has a high income and does not receive transfers, she will be inclined to exaggerate when asked about the desirability of public good provision. Likewise, an individual with a low level of income tends to understate the desirability of public good provision because he fears a reduction of income transfers.

The difficulty in finding an optimal mechanism for both redistribution and public good provision is that there are two incentive problems simultaneously. The first one is familiar from the theory of optimal income taxation and is due to the fact that individuals have private information on their earning abilities. This imposes incentive constraints on redistribution which give rise to what is known as the *equity-efficiency tradeoff*.[1] The second problem is the classical *free-rider problem*, which arises because individuals have private information on their valuation of a non-excludable public good. The main insight from the joint analysis of these two incentive problems is that the *equity-efficiency tradeoff* and the *free-rider problem* interact in a systematic fashion. More able individuals can have an excessive desire for public good provision, which they value as an instrument to limit the extent of redistribution. Likewise, less able individuals may tend to understate the desirability of provision in order to avoid a cut of transfers. Hence, a decision on provision that reflects the "true" aggregate valuation of the public good necessitates an adjustment of the transfer system that corrects these biases. This requires a *complementarity* between the level of redistribution and the decision on public good provision, relative to an *equity-efficiency tradeoff* without a *free-rider problem*: To prevent the more productive class from exaggerating, public good provision has to be accompanied by an increased level of redistribution. Similarly, the less productive are prevented from understating their valuation of the public good by a reduced level of redistribution if there is no public good provision.

[1]This literature starts with Mirrlees (1971). See Hellwig (2005) for a recent treatment.

The model that is used to arrive at these results combines a *screening problem* with a problem of *information aggregation* and involves two dimensions of individual heterogeneity, earning abilities as well as preferences for the public good. More precisely, the following assumptions are made. Individuals either have a low or a high level of earning ability.[2] Likewise, valuations of the public good are either high or low. Moreover, public goods preferences are assumed to be perfectly correlated with earning ability. That is, all individuals with the same level of earning ability also have the same valuation of the public good. With this specific information structure, the *screening problem* is to identify which individual has been assigned which level of earning ability. The problem of *information aggregation* is the elicitation of the public goods preferences of high and low ability individuals, respectively.[3]
As is standard in the literature on optimal income taxation, the present paper assumes that there is a continuum of agents. While this assumption has a variety of convenient implications, it creates a difficulty when trying to discuss problems of *information aggregation* under incentive constraints. One might argue that, in a large economy, *free-rider problems* do not arise as a single individual has no impact on public good provision and hence no reason to hide his true valuation. However, the present paper takes a different view, based on the observation that, in a continuum economy, collective behavior of individuals has an impact on the perceived aggregate valuation of the public good. Indeed as will be shown below, allocation rules based on income tax schedules are vulnerable to coordinated manipulations by large groups of agents. The notion of a *collectively incentive compatible* income tax is introduced to deal with this issue. It specifies collective incentive conditions that ensure that *information aggregation* may proceed even under the threat of manipulative collective behavior.[4]

The remainder of the paper is organized as follows. Section 2 defines the environment. As a benchmark, Section 3 derives the optimal income taxation without a *free-rider problem*. Section 4 contains the definition of a *collectively incentive compatible income tax*. In section 5 the *optimal* collectively incentive compatible income tax is characterized. The last section contains concluding remarks. All proofs can be found in the appendix.

[2]This two-class economy is a special case that has received some attention in the literature on optimal taxation. See e.g. Mirrlees (1975), Stiglitz (1982, 1987), Boadway and Keen (1993), Nava et al. (1996) or Gaube (2005).

[3]Consequently, the *screening* problem is based on only one dimension of individual heterogeneity. There cannot be a discrimination between individuals with the same earning ability but different public goods preferences, as in Hellwig (2004). This author however assumes that there is no problem of *information aggregation*.

[4]This solution concept has been inspired by the literature on mechanism design problems under a threat of collusion among agents, most notably Bernheim and Whinston (1986), Laffont and Martimort (1997, 1999) and Demange and Guesnerie (2001).

2 The environment

The economy consists of a continuum of individuals $j \in I := [0,1]$. An individual has a pair of characteristics (w^j, θ^j), where w^j is a productivity parameter and θ^j is a taste parameter for a public good. w^j and θ^j are taken to be the realizations of the binary random variables $\tilde{w}^j$ and $\tilde{\theta}^j$, respectively. The possible values w_1, w_2 of $\tilde{w}^j$ and θ_L, θ_H of $\tilde{\theta}^j$ are taken to be the same for all j. Without loss of generality, $w_1 < w_2$ and $\theta_L < \theta_H$.

The random variables $\tilde{w}^j$, $j \in I$, are assumed to satisfy a *Law of Large Numbers for large economies*:[5] while each individual has probability 1/2 for a high or a low productivity realization, this uncertainty about productivity parameters disappears in the aggregate. Ex post, after the realization of individual uncertainty, there are equal shares of more and less productive individuals in the population. For brevity, I refer to those individuals, who end up with the low productivity parameter w_1, as class 1 individuals. Likewise, the individuals with productivity parameter w_2 are called class 2 individuals.

The random variables $\tilde{\theta}^j$, $j \in I$, are assumed to be perfectly correlated among all individuals with the same productivity parameter, i.e. ex post all individuals of class t, $t \in \{1,2\}$, have the same taste parameter. Let θ_t be the common value of the taste parameter $\tilde{\theta}^j$ for all individuals j with $\tilde{w}^j = w_t$.

The taste parameters θ_1 and θ_2 are the realizations of random variables $\tilde{\theta}_1$ and $\tilde{\theta}_2$. The economy as a whole is subject to uncertainty about these random variables. There are four possibilities, or *states*, denoted by s_{LL} , s_{LH} ,s_{HL} and s_{HH}, where, e.g. s_{LL} indicates that $\tilde{\theta}_1 = \theta_L$ and $\tilde{\theta}_2 = \theta_L$. Analogously, s_{LH} indicates that $\tilde{\theta}_1 = \theta_L$ and $\tilde{\theta}_2 = \theta_H$, etc. The set of states is written as $S = \{s_{LL}, s_{LH}, s_{HL}, s_{HH}\}$.

All individuals of type t have the same utility function, which takes the form

$$U_t \quad = \quad \theta_t Q + u(C) - v\left(\frac{Y}{w_t}\right) \; . \tag{1}$$

C denotes consumption of private goods and $Y = Lw_t$ denotes effective labor or income. That is, w_t can be interpreted as a wage rate and L denotes hours worked to generate income Y. Obviously, to achieve a given income Y individuals with a lower wage have to work more. $Q \in \{0,1\}$ stands for a public project, which is either installed or not. The functions u and v are strictly increasing and twice continuously differentiable. Moreover, u is concave and v is convex. In addition, those functions satisfy the following boundary condition, which ensures interior solutions to optimization problems: for all w_t and all $C > 0$, there exists $Y > 0$, such that

$$u'(C) - \frac{1}{w_t} v'\left(\frac{Y}{w_t}\right) \quad = \quad 0 \; .$$

[5]For a formal discussion, see Judd (1985) or Al-Najjar (2004).

Finally, note that preferences satisfy the single crossing condition with respect to the productivity parameter. Accordingly, at any point in the Y-C plane, the indifference curve of a less productive individual is steeper.

Information Structures

Throughout the analysis, I assume that the parameter values w_1, w_2, θ_L and θ_H are common knowledge. In contrast, the assignment of any one individual to the more or less productive class is that individual's private information. This privacy of information gives rise to *assignment uncertainty*.

Further, I distinguish between two model specifications according to whether the realizations of $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are common knowledge. The model has *pure assignment uncertainty* if these realizations, and hence the state of the world $s \in S$, are commonly known. The model exhibits *private information on taste parameters* if only individuals of class t observe wether $\tilde{\theta}_t = \theta_L$ or $\tilde{\theta}_t = \theta_H$. In the latter case, in addition to the uncertainty regarding individuals' class assignments, there is aggregate uncertainty with respect to *unknown class characteristics.*

Anonymous Allocations and Income Tax Mechanisms

The analysis of admissible allocations is treated as a problem of mechanism design. Attention is restricted to the class of anonymous allocation mechanisms which are *individually incentive compatible* and feasible. In particular, this class of allocation mechanisms is flexible enough to deal with both information structures.

An *anonymous allocation mechanism* specifies for each state $s \in S$ a public good provision level $Q(s)$ and for each characteristic in $(w, \theta) \in \Gamma := \{w_1, w_2\} \times \{\theta_L, \theta_H\}$ a consumption level $C(w, \theta, s)$ and an output requirement $Y(w, \theta, s)$. An anonymous allocation mechanism is *individually incentive compatible (I-IC)* if $\forall s \in S$, $\forall (w, \theta) \in \Gamma$ and $\forall (\hat{w}, \hat{\theta}) \in \Gamma$,

$$u(C(w,\theta,s)) - v\left(\frac{Y(w,\theta,s)}{w}\right) \quad \geq \quad u(C(\hat{w},\hat{\theta},s)) - v\left(\frac{Y(\hat{w},\hat{\theta},s)}{w}\right) \; .$$

An anonymous allocation mechanism is *feasible* if $\forall s \in S$, $\forall (w, \theta) \in \Gamma_r(s)$,

$$Y(w_1,\theta_1,s) - C(w_1,\theta_1,s) + Y(w_2,\theta_2,s) - C(w_2,\theta_2,s) \geq kQ(s) \; ,$$

where k denotes the cost of public good provision and $\Gamma_r(s)$ the set of individual characteristics supported in state s, e.g. $\Gamma_r(s_{LH}) = \{(w_1, \theta_L), (w_2, \theta_H)\}$.

Some explanatory remarks are in order. The *I-IC* constraints specify incentives on the individual level. As the economy is large, those constraints are stated for a given state s. This reflects the fact that, in a large economy, no single individual is able to influence the state of the world as perceived by the mechanism designer. In particular, no individual has a noticeable

impact on public good provision.
If the information structure exhibits private information on taste parameters, then the tax setting institution has to deduce the actual state from individual reports. That is, the mechanism designer receives from each individual a statement which consists of an announced earning ability level and an announced taste parameter. An evaluation of all individual reports makes it possible to observe whether the less (more) able individuals have a low or a high taste parameter. The fact that s can not be taken as given explains why the message set in the revelation game equals Γ. However, if the analysis is concerned with pure assignment uncertainty, the message set $\Gamma_r(s)$ is sufficient.[6]
The *I-IC* conditions require that truth-telling constitutes an equilibrium in weakly dominant strategies.[7] That is, truth-telling has to be a best-response from an individual's perspective, irrespective of the announcements of others and irrespective of the actual state of the world. However, the *I-IC* conditions specify individual incentives only in response to message profiles that indicate a feasible state of the economy. A complete description of the revelation game also requires a specification of what happens if this distribution is incompatible with what is commonly known about the set S. These out-off-equilibrium payoffs have to preserve the incentive structure, i.e. they have to be such that truth-telling is a weakly dominant strategy. This is, for instance, achieved by choosing $Q = 0$ and a degenerate consumption-income menu that contains only one C-Y-combination.
The final remark clarifies why the set of anonymous, feasible and *I-IC* allocation mechanisms is of relevance for an analysis of income tax systems. To this end, call an anonymous allocation mechanism an *income tax mechanism* if there exists a function $T : \mathbb{R}_+ \times S \rightarrow \mathbb{R}$ such that $\forall (w, \theta) \in \Gamma, \forall s \in S$:

i) $C(w, \theta, s) = Y(w, \theta, s) - T(Y(w, \theta, s), s)$

ii) $Y(w, \theta, s) \in \text{argmax}_Y u(Y - T(Y, s)) - v\left(\frac{Y}{w}\right)$.

and, moreover, such that $\forall s \in S$ and $(w_1, \theta_1), (w_2, \theta_2) \in \Gamma_r(s)$,

$$T(Y(w_1, \theta_1, s)) + T(Y(w_2, \theta_2, s)) \geq kQ(s) \ .$$

As has been shown by Hammond (1979) and Guesnerie (1995), the set of *income tax mechanisms* can be equivalently analyzed via the set of *I-IC* and *feasible* allocation mechanisms. Formally, one has the following result: An anonymous allocation mechanism is *I-IC* and feasible if and only if it is an *income tax mechanism.*

[6]The revelation principle implies that any further element of the message set would be superfluous.

[7]The advantage of *implementation in dominant strategies* – relative to other solution concepts – is that individual behavior neither depends on a *common prior assumption* nor on a specific form of strategic reasoning in case of multiple equilibria.

The following lemma provides an alternative characterization of income tax mechanisms which proves helpful in subsequent sections.

Lemma 1 An anonymous allocation mechanism is an income tax mechanism if and only if it is feasible and possesses the following properties:

i) *No discrimination of taste in terms of utility (NDT-U)*: $\forall s \in S$, $\forall w \in \{w_1, w_2\}$, $\forall \theta \in \{\theta_L, \theta_H\}$ and $\forall \theta' \in \{\theta_L, \theta_H\}$,

$$u(C(w, \theta, s)) - v\left(\frac{Y(w, \theta, s)}{w}\right) = u(C(w, \theta', s)) - v\left(\frac{Y(w, \theta', s)}{w}\right) .$$

ii) *Individual revelation of productivity (I-RP)*: $\forall s \in S$, $\forall \theta \in \{\theta_L, \theta_H\}$, $\forall t \in \{1, 2\}$ and $t \neq t'$,

$$u(C(w_t, \theta, s)) - v\left(\frac{Y(w_t, \theta, s)}{w_t}\right) \geq u(C(w_{t'}, \theta, s)) - v\left(\frac{Y(w_{t'}, \theta, s)}{w_t}\right) .$$

The lemma follows from the fact that individuals take the state s and hence the level of public good provision as given. Due to the additive separability of preferences, this implies that individual incentive conditions become independent of taste parameters. Consequently, an income tax mechanism can use only individual differences in productivity as a screening device.

3 Pure assignment uncertainty

Contributions to the theory of optimal utilitarian income taxation are typically concerned with the case of pure assignment uncertainty. This section recalls results from this literature for the special setup of a two-class economy and derives a further comparative statics property. This provides a benchmark case, that proves helpful for the analysis of an information structure with private information on taste parameters in later sections.

3.1 The optimization problem

Under pure assignment uncertainty the state s of the economy is commonly known. Equivalently, for each taste parameter $\tilde{\theta}_t$, $t \in \{1, 2\}$, it is commonly known whether the realization θ_t equals θ_L or θ_H. Consequently, assignment uncertainty stems only from the fact that each individual i has private information on whether her productivity parameter equals w_1 or w_2. This considerably simplifies the analysis of anonymous allocation mechanisms. Once individual productivity is revealed, an individual's class assignment is known, and so is the individual's taste parameter. Hence, there is no need to specify C-Y pairs that depend on declared taste parameters. For the remainder of this section, I may thus suppress the dependence on taste parameters and write $C_t(s)$ and $Y_t(s)$ instead of $C(w_t, \theta_t, s)$ and $Y(w_t, \theta_t, s)$. Under pure assignment uncertainty, an income tax mechanism is a collection $\{Q(s), Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s \in S}$, which satisfies, for all s, the feasibility constraints

$$Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq kQ(s), \quad Q(s) \in \{0, 1\} \ , \tag{2}$$

and the *I-RP* constraints

$$\begin{aligned} u(C_1(s)) - v\left(\frac{Y_1(s)}{w_1}\right) &\geq u(C_2(s)) - v\left(\frac{Y_2(s)}{w_1}\right) \ , \\ u(C_2(s)) - v\left(\frac{Y_2(s)}{w_2}\right) &\geq u(C_1(s)) - v\left(\frac{Y_1(s)}{w_2}\right) \ . \end{aligned} \tag{3}$$

Note that, under pure assignment uncertainty, the *NDT-U* property is moot. There is no need to specify a *C-Y* pair for individuals who claim a "wrong" taste parameter. The "true" taste parameter is known anyway once an individual's productivity level is revealed.

In state s, an income tax mechanism generates a utilitarian welfare level, which is, in the following, written as

$$W(s) := (\theta_1 + \theta_2)Q(s) + u(C_1(s)) - v\left(\frac{Y_1(s)}{w_1}\right) + u(C_2(s)) - v\left(\frac{Y_2(s)}{w_2}\right) \ .$$

Under pure assignment uncertainty, the state s is commonly known. Hence, it might seem natural to define an *optimal* utilitarian income tax mechanism such that, for given s, $W(s)$ is maximized subject to the feasibility constraints in (2) and the *I-RP* constraints in (3). I will, however, proceed differently. Below a definition is stated which yields trivially the same set of optimal allocations, but facilitates a comparison to the case of private information on taste parameters discussed in later sections.

An income tax mechanism is evaluated from an ex ante perspective, which is defined as a hypothetical situation where the actual state s is not yet known. That is, the objective function is a weighted average of the welfare levels in $\{W(s)\}_{s \in S}$, with a probability weight attached to each state s. These probability weights are taken to be the prior beliefs of the tax setting planner who perceives the actual state s of the economy as the realization of a random variable $\tilde{s}$. The prior beliefs are denoted $p := (p_{LL}, p_{LH}, p_{HL}, p_{HH})$, where $p_{LL} := \text{prob}(\tilde{s} = s_{LL})$, $p_{LH} := \text{prob}(\tilde{s} = s_{LH})$, etc. Expected welfare from the planner's ex ante perspective is accordingly given by

$$EW \ := \ p_{LL}W(s_{LL}) + p_{LH}W(s_{LH}) + p_{HL}W(s_{HL}) + p_{HH}W(s_{HH}) \ .$$

Definition 1 Under pure assignment uncertainty, an optimal income tax mechanism solves the problem of choosing $\{Q(s), Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s \in S}$ in order to maximize EW subject to the feasibility constraints in (2) and the *I-RP* constraints in (3).

For brevity, I refer to this optimal income tax problem under pure assignment uncertainty as the *informed problem* and to its solution as the *informed optimum.*

Characterizing the informed optimum

For a characterization of the informed optimum, it is helpful to introduce the following auxiliary problem, which does not include a public good but instead an exogenous revenue requirement $r \geq 0$ in the budget constraint.

$$\begin{aligned} &\max_{C_1,Y_1,C_2,Y_2} && u(C_1) - v\left(\frac{Y_1}{w_1}\right) + u(C_2) - v\left(\frac{Y_2}{w_2}\right) \\ &\text{s.t.} && Y_1 - C_1 + Y_2 - C_2 \geq r \ , \\ &&& u(C_1) - v\left(\frac{Y_1}{w_1}\right) \geq u(C_2) - v\left(\frac{Y_2}{w_1}\right) \ , \\ &&& u(C_2) - v\left(\frac{Y_2}{w_2}\right) \geq u(C_1) - v\left(\frac{Y_1}{w_2}\right) \ . \end{aligned} \tag{4}$$

A solution to problem (4) is parameterized by the revenue requirement r and denoted $(Y_1^*(r), C_1^*(r), Y_2^*(r), C_2^*(r))$. The following result is well known (see e.g. Stiglitz (1982)).

Lemma 2 At a solution to problem (4) the feasibility constraint and only the *I-RP* constraint for $t = 2$ are binding, implying that there is a *distortion at the bottom* and *no distortion at the top*:

$$MRS_1^* := \frac{\frac{1}{w_1} v'\left(\frac{Y_1^*(r)}{w_1}\right)}{u'(C_1^*(r))} < 1 \quad \text{and} \quad MRS_2^* := \frac{\frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right)}{u'(C_2^*(r))} = 1 \ .$$

Intuitively, problem (4) is essentially a problem of redistribution under incentive constraints. As the more productive suffer less from the necessity to generate income, a utilitarian planner wants them to work harder. This implies a binding *I-RP* constraint for this class of individuals at the informed optimum.

The informed optimum is now characterized with reference to problem (4). I use a shorthand notation for the utility level at a solution to problem (4) induces for type t individuals:[8]

$$R_t(r) \ := \ u(C_t^*(r)) - v\left(\frac{Y_t^*(r)}{w_t}\right) \ .$$

[8]I use the letter R to indicate that I refer to a utility level which is generated by a solution to an optimization problem with an exogenous *R*evenue *R*equirement.

Obviously, the informed utilitarian planner decides on public good provision according to the following criterion: $Q(s) = 1$ if and only if

$$\theta_1 + \theta_2 \quad \geq \quad R_1(0) + R_2(0) - \Big(R_1(k) + R_2(k)\Big) \ .$$

Under this criterion, the provision rule chosen by an informed planner depends on the parameter values θ_L and θ_H. E.g. if

$$2\theta_L \quad > \quad R_1(0) + R_2(0) - \Big(R_1(k) + R_2(k)\Big) \ ,$$

then an informed planner chooses $Q(s) = 1$ for all s. To avoid a lengthy discussion of each conceivable parameter constellation, I focus on a particular case.

Assumption 1 An informed planner chooses to install the public good in all states except state s_{LL}:[9]

$$\theta_H + \theta_L \quad \geq \quad R_1(0) + R_2(0) - \Big(R_1(k) + R_2(k)\Big) \quad \geq \quad 2\theta_L \ .$$

For ease of reference, I denote by $Q^i : Q = 0 \Longleftrightarrow s = s_{LL}$ the provision rule chosen by an informed planner. To complete the description of the informed optimum, I denote by $U_1^i(s)$ and $U_2^i(s)$ the realized utility levels of class 1 and class 2 individuals. Obviously,

$$U_1^i(s) \quad = \quad \begin{cases} R_1(0), & \text{if } s = s_{LL}, \\ \theta_L + R_1(k), & \text{if } s = s_{LH}, \\ \theta_H + R_1(k), & \text{if } s = s_{HL}, \\ \theta_H + R_1(k), & \text{if } s = s_{HH} \end{cases} \qquad \text{and}$$

$$U_2^i(s) \quad = \quad \begin{cases} R_2(0), & \text{if } s = s_{LL}, \\ \theta_H + R_2(k), & \text{if } s = s_{LH}, \\ \theta_L + R_2(k), & \text{if } s = s_{HL}, \\ \theta_H + R_2(k), & \text{if } s = s_{HH}. \end{cases}$$

I refer to the expression $R_t(0) - R_t(k)$ as the *utility loss* of class t from paying for public good provision at the informed optimum. Moreover, I say that for class t individuals, the *willingness to pay for the public good* is positive (negative) if the *utility gain* θ_t exceeds (falls short of) this utility loss, i.e. if $\theta_t - (R_t(0) - R_t(k))$ is positive (negative).

[9] Obviously, a parameter constellation such that $Q = 1$ is desired in every (no) state of the world is not very interesting. Hence, the only alternative of interest is that $Q = 0$ is preferred in states s_{LH} and s_{HL}. An investigation of this case gives rise to an analysis which is analogous to the one presented below.

3.2 Conflicting interests at the informed optimum

Even though an optimal utilitarian income tax attaches equal weight to the utility levels realized by the more and the less able class of individuals, the informed optimum may give rise to conflicting views on the desirability of public good provision. To illustrate this, suppose for the sake of concreteness that

$$R_1(0) - R_1(k) \quad > \quad \theta_H \quad > \quad \theta_L \quad > \quad R_2(0) - R_2(k) \; . \tag{5}$$

In this scenario, for the more productive individuals, the utility loss is so small that their willingness to pay for the public good is positive in all states s. By contrast, the less productive suffer so severely from the increased revenue requirement if the public good is installed that they oppose provision in every state of the world.

A clarification of the possible patterns of conflicting interests will be important for an understanding of the additional incentive problems that come into play under an information structure with private information on taste parameters. Intuitively, if the scenario characterized by the inequalities in (5) arises, less productive individuals want to prevent the public good from being installed in every state s, and hence they have an incentive to report a low taste realization even if in fact their taste parameter is high. Likewise, the more able class wants to get the public good in every state and might be tempted to report a high taste in case of a low taste realization.

The following lemma is important for an understanding of possible scenarios of conflicting interests. It shows that for the less productive class of individuals the utility loss is larger if in problem (4) the revenue requirement r is increased. In more technical terms, the lemma establishes a property of *decreasing differences* according to which a lower productivity level translates into a larger utility loss. The proof relies on the following assumption:

Assumption 2 The function v is strictly convex and satisfies[10]

$$\forall x \geq 0: \quad \frac{1}{w_1^2} v''\left(\frac{x}{w_1}\right) \quad \geq \quad \frac{1}{w_2^2} v''\left(\frac{x}{w_2}\right) \; .$$

Lemma 3 Let $v(\cdot)$ satisfy Assumption 2. Let $r' > r$. Then:

$$R_1(r) - R_1(r') \quad > \quad R_2(r) - R_2(r') \quad > \quad 0 \; .$$

The intuition behind this observation is as follows: Consider a solution to problem (4) and suppose the revenue requirement is slightly increased.

[10] Note that a sufficient condition for Assumption 2 is $v''' \geq 0$. An alternative assumption, which would also yield the result of Lemma 3, is that the function v is linear. For a discussion of this quasi-linear case, see Weymark (1986) or Boadway et al. (2000).

The more productive cannot be forced to cover the resulting small budget deficit, as this would violate their *I-RP* constraint. To the contrary, less able individuals can be made worse off without violating any constraint. Consequently, the planner has to make them worse off if there is a need to extract larger revenues.

Possible scenarios of conflicting interests

If combined with the observation that the utility loss is larger for less able individuals, as shown in lemma 3, assumption 1 implies that the willingness of less able individuals to pay is negative if $\theta_1 = \theta_L$. Analogously, for the more productive class, the willingness to pay is positive if $\theta_2 = \theta_H$,

$$R_1(0) - R_1(k) > \theta_L \quad \text{and} \quad \theta_H > R_2(0) - R_2(k) \; . \tag{6}$$

These inequalities in conjunction with assumption 1 reduce the set of possible parameter constellations. The following three scenarios may arise.

$$\text{Sc.1:} \quad \theta_H \;\geq\; R_1(0) - R_1(k) \;>\; R_2(0) - R_2(k) \;\geq\; \theta_L \; ,$$

$$\text{Sc.2:} \quad \theta_H \;\geq\; R_1(0) - R_1(k) \;\geq\; \theta_L \;>\; R_2(0) - R_2(k) \; ,$$

$$\text{Sc.3:} \quad R_1(0) - R_1(k) \;>\; \theta_H \;>\; \theta_L \;>\; R_2(0) - R_2(k) \; .$$

These inequalities are interpreted as follows.
Scenario 1: For individuals of any class t, willingness to pay for the public good is positive if the taste realization is high, $\tilde{\theta}_t = \theta_H$, and is negative if the taste realization is low, $\tilde{\theta}_t = \theta_L$. Scenario 1 hence gives rise to the statement that, at the *informed optimum*, *willingness to pay for the public good is independent of earning ability.*
Scenario 2: For the less productive class, as under *Scenario 1*, the willingness to pay for the public good is positive only if the utility gain is high. In contrast, more productive individuals, whose utility loss is smaller, have a positive willingness to pay in any state s.
Scenario 3: For more productive individuals, as under *Scenario 2*, the willingness to pay for the public good is always positive. In addition, less able individuals suffer from such a heavy utility loss that their willingness to pay is negative in any state s.

4 Private information on taste parameters

From now on, I consider an information structure with private information on taste parameters. Consequently, a utilitarian planner faces the problem of information aggregation simultaneously with the screening problem of identifying which individual belongs to which class. This necessity of information aggregation will in general cause additional incentive problems, on

top of the *I-RP* requirement.
To illustrate this, suppose *Scenario 2* applies and ask whether the informed optimum is implementable. If one takes the view that individual incentives are enough, the answer is yes. As all individuals take the mechanism designer's perception of the actual state as outside their influence, no isolated individual has a reason to misreport her own taste parameter. However, if the informed optimum is implemented, the more productive individuals want to have the public good in all states of the world, that is, even if $\tilde{\theta}_2 = \theta_L$. And moreover, if class 2 individuals are able to convince the utilitarian planner that their taste parameter is in fact high, they can ensure the provision of the public good. As the decision on provision is based on a revelation game, there is an obvious way to achieve this: a collective lie of all class 2 individuals on their taste parameter.
These considerations highlight the following issues: *First*, with private information on taste parameters, a mechanism designer may not be able to detect a deviation from the truth by a subset of agents. If all class 2 individuals make the same announcement $\hat{\theta}_2$, it is not possible to tell whether those individuals are jointly lying or are jointly telling the truth. *Second*, such a deviation may be beneficial for such a subset of agents. *Third*, it is not prevented by individual incentive compatibility. Given that all class 2 individuals lie about their taste parameter, there is no incentive for an isolated class 2 individual to reveal the realization of $\tilde{\theta}_2$ truthfully. Due to the *NDT-U* property of income tax mechanisms, this is a systematic feature. A collective deviation involving taste parameters is not undermined by individual incentives.

4.1 Collective Incentive Compatibility

In the following a *collectively incentive compatible (C-IC)* income tax mechanism is defined. Such a mechanism ensures that truth-telling is an equilibrium outcome even under the threat of collective manipulations.
Denote by $\mathcal{J}$ the set of measurable subsets of the set of agents, $I = [0, 1]$, with positive length. A typical element is denoted J. Denote the true profile of characteristics in J by $\gamma_J := \{(w^j, \theta^j)\}_{j \in J}$. Denote the reported profile by $\hat{\gamma}_J := \{(\hat{w}^j, \hat{\theta}^j)\}_{j \in J}$.
Denote the cross-section distribution of announcements induced by $\hat{\gamma}_J$ if the true state of the economy is $s \in S$ and all individuals not in J report truthfully by $\delta(\hat{\gamma}_J, s)$. Note that any such distribution belongs to the set $\Delta(\Gamma)$ of probability distributions on $\Gamma = \{w_1, w_2\} \times \{\theta_L, \theta_H\}$, i.e. it assigns a probability weight to each of the four elements of Γ.
Denote by $\mathcal{D} := \{d_{LL}, d_{LH}, d_{HL}, d_{HH}\}$ the set of cross-section distributions of characteristics which correspond in an obvious way to the possible states of the world, e.g. d_{LH} is a distribution which assigns equal mass to the elements of $\Gamma_r(s_{LH}) = \{(w_1, \theta_L), (w_2, \theta_H)\}$. For $\delta(\hat{\gamma}_J, s) \in \mathcal{D}$, denote by

$\hat{s}(\hat{\gamma}_J, s) \in S$, the *perceived* state of the world, e.g. if $\delta(\hat{\gamma}_J, s) = d_{LH}$, then $\hat{s}(\hat{\gamma}_J, s) = s_{LH}$.

Definition 2 A coalition J is said to *manipulate* an income tax mechanism if there exists $s \in S$ and $\hat{\gamma}_J \neq \gamma_J$ with the following properties:

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

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

$$\theta^j Q(\hat{s}(\hat{\gamma}_J, s)) + u(C(\hat{w}^j, \hat{\theta}^j, \hat{s}(\hat{\gamma}_J, s))) - v\left(\frac{Y(\hat{w}^j, \hat{\theta}^j, \hat{s}(\hat{\gamma}_J, s))}{w^j}\right)$$

$$> \theta^j Q(s) + u(C(w^j, \theta^j, s)) - v\left(\frac{Y(w^j, \theta^j, s)}{w^j}\right) \ .$$

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

$$\theta^j Q(\hat{s}(\hat{\gamma}_J, s)) + u(C(\hat{w}^j, \hat{\theta}^j, \hat{s}(\hat{\gamma}_J, s))) - v\left(\frac{Y(\hat{w}^j, \hat{\theta}^j, \hat{s}(\hat{\gamma}_J, s))}{w^j}\right)$$

$$= \theta^j Q(\hat{s}(\hat{\gamma}_J, s)) + u(C(w^j, \theta^j, \hat{s}(\hat{\gamma}_J, s))) - v\left(\frac{Y(w^j, \theta^j, \hat{s}(\hat{\gamma}_J, s))}{w^j}\right) \ .$$

iv) *Collective Stability.* There does not exist a subcoalition $K \subset J$, with an *undetectable* collective deviation $\tilde{\gamma}_K \neq \hat{\gamma}_K$ that induces a state perception $\hat{s}(\tilde{\gamma}_K, \hat{\gamma}_{J\setminus K}, s)$ that makes all members of K strictly better off relative to $\hat{s}(\hat{\gamma}_J, s)$ (*unanimity*), prescribes for all its members individually best responses given the state perception $\hat{s}(\tilde{\gamma}_K, \hat{\gamma}_{J\setminus K}, s)$ (*individual stability*) and is not threatened by further collective manipulations, which satisfy all these requirements (*collective stability*).

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

According to this definition, a coalition considers a collective deviation in response to truth-telling of all other individuals. The scope for manipulation is limited by the requirement that it must not be *detectable*, i.e. the relevant coalitional plans are only those for which it does not become apparent that a manipulation has occurred. Moreover, coalition members have to agree unanimously on a deviation and may not use side payments to reach such an agreement. Finally, a coalition has to meet two *stability* requirements. The incentives coalition members face individually must not conflict with the message profile used by the coalition; that is, collective manipulations are a concern only in so far as they do not conflict with *I-IC*. In addition, a conceivable collective manipulation must not provoke the formation of a subcoalition which departs from the original coalitional plan. These stability requirements have been introduced by Bernheim et al. (1986) in their definition of a *coalition-proof* Nash-equilibrium.

A peculiarity of Definition 2 is that the collective stability of a coalition J is defined with reference to the collective stability of a coalition $K \subset J$. Obviously, in a continuum economy, there is no chance of tracing these notions back to the collective stability of some "smallest" coalitions. As will become clear (see Proposition 1), for the purposes of this paper, this does not create a problem. The structure of a two-class economy is sufficiently simple to arrive at a complete characterization of *C-IC* income tax mechanisms.

With reference to the literature, different interpretations of the implicit assumptions on coalition formation can be given. First, suppose that pre-play communication resolves the uncertainty among individuals about the actual state of the economy.[11] The above definition then requires that truth-telling is a best response from the perspective of a coalition whose members know the true state of the world and presume that all individuals outside the coalition tell the truth. Alternatively, *C-IC* can be framed as a *robustness*-requirement.[12] It implies that ex post, after the state of the world has become commonly known, no subset of individuals would *jointly* want to revise their announcements if they were, hypothetically, given the opportunity to do so.

4.2 *C-IC* in the two-class economy

The definition of *C-IC* stated above is rather abstract in the sense that it excludes any kind of coalitional manipulation. This concern can be simplified by making use of the specific features of a two-class economy. As developed below, it suffices to exclude manipulative threats of coalitions, which consist of all individuals of one class. Moreover, individual and collective incentive concerns can be separated: the latter require that individuals belonging to the same class are prevented from a collective lie on their taste parameter, while the former ensure a revelation of productivity parameters.

Definition 3 A *utility allocation* specifies for every state $s \in S$, utility levels $\tilde{U}_1(s)$ and $\tilde{U}_2(s)$ for type 1 and type 2 individuals, respectively. A

[11] Such pre-play communication works if one assumes that individuals are able to solve pure coordination problems by cheap talk, Farrell and Rabin (1996).

[12] *Robustness* requires that the set of implementable allocations does not depend on assumptions about the prior beliefs of individuals. For a more extensive discussion, see, e.g. Bergemann and Morris (2005); Chung and Ely (2004) or Kalai (2004).

utility allocation is said to be *implementable* if there exists a *C-IC* income tax mechanism such that $\forall s \in S$, and for all $(w_t, \theta_t) \in \Gamma_r(s)$,

$$\begin{aligned}\tilde{U}_t(s) &= \theta_t Q(s) + u(C(w_t, \theta_t, s)) - v\left(\frac{Y(w_t, \theta_t, s)}{w}\right) \\ &=: \theta_t Q(s) + V_t(s) \ ,\end{aligned}$$

where $V_t(s)$ is a shorthand for the utility class t individuals derive in state s from their consumption-income combination.
A utility allocation $\{\tilde{U}_1(s), \tilde{U}_2(s)\}_{s \in S}$ is said to be *Pareto-optimal* if it is implementable and there does not exist some other implementable utility allocation $\{\tilde{U}'_1(s), \tilde{U}'_2(s)\}_{s \in S}$ which yields, in all states s and for all $t \in \{1, 2\}$, a weakly larger utility level, $\tilde{U}'_t(s) \geq \tilde{U}_t(s)$, and in some state s and for some class t a strictly larger utility level, $\tilde{U}'_t(s) > \tilde{U}_t(s)$.

Proposition 1 Suppose there is no pooling of earning ability, that is, $\forall s \in S, \forall (w_t, \theta_t) \in \Gamma_r(s)$, $(C(w_1, \theta_1, s), Y(w_1, \theta_1, s)) \neq (C(w_2, \theta_2, s), Y(w_2, \theta_2, s))$.[13] Then, a utility allocation is Pareto-optimal if and only if it is implementable by a feasible allocation mechanism which satisfies *I-RP* and the following properties:

i) *Collective revelation of taste on the class level (C-RT-C)*: $\forall x \in \{L, H\}$, $\forall \hat{x} \in \{L, H\}$, $\forall y \in \{L, H\}$ and $\forall \hat{y} \in \{L, H\}$:

$$\theta_x Q(s_{xy}) + V_1(s_{xy}) \geq \theta_x Q(s_{\hat{x}y}) + V_1(s_{\hat{x}y}) \ ,$$

$$\theta_y Q(s_{xy}) + V_2(s_{xy}) \geq \theta_y Q(s_{x\hat{y}}) + V_2(s_{x\hat{y}}) \ .$$

ii) *No discrimination of taste in terms of consumption and income (NDT-CY)*: $\forall s \in S$, $\forall w \in \{w_1, w_2\}$, $\forall \theta \in \{\theta_L, \theta_H\}$ and $\forall \theta' \in \{\theta_L, \theta_H\}$,

$$(C(w, \theta, s), Y(w, \theta, s)) = (C(w, \theta', s), Y(w, \theta', s)) \ .$$

The *NDT-CY* property requires that, for a given distribution of characteristics in the economy, the allocation of private goods is independent of taste parameters. This is a slightly stronger property as relative to *NDT-U*. According to the *C-RT-C*-property, manipulations of coalitions consisting only of individuals with the same type and which misreport only the taste parameter are ruled out. Obviously, this condition is necessary for *C-IC*. Proposition 1 states that it is also sufficient if one restricts attention to Pareto-optimal allocations.

[13] Absence of pooling is required only to make the presentation more accessible. In subsequent sections, optimal tax mechanisms are characterized without imposing this assumption. It will turn out that an optimum does not involve pooling.

The proof proceeds as follows. First it is shown that there cannot be an undetectable collective manipulation that involves productivity parameters. This would require some type 1 individuals to be willing to claim a high productivity and some type 2 individuals to be willing to claim a low productivity. Due to the single-crossing property, this is not compatible with *I-IC* unless there is pooling. Then, it is observed that undetectability in a two-class economy requires all individuals who report the same productivity parameter to agree on the reported taste parameter as well. Hence, there remain only two kinds of collective manipulations: those where only the individuals of one class lie on their taste parameter and those where individuals of both classes jointly lie on their taste parameter. The former kind of collective manipulation is ruled out by the *C-RT-C* property. The latter would require that both classes prefer a different state perception. It is shown that this situation can not arise under a Pareto-optimal utility allocation.

Proposition 1 justifies the restriction to allocation rules with the *NDT-CY* property. This implies that a more concise notation can be used. In the following, the consumption and income for an individual of type t, given that the state of the world is s, is written as $(C_t(s), Y_t(s))$, with the understanding that this pair equals both $(C(w_t, \theta_L, s), Y(w_t, \theta_L, s))$ and $(C(w_t, \theta_H, s), Y(w_t, \theta_H, s))$. The *I-RP* property is hence written in the following as $\forall s \in S$, $\forall t$, $\forall t' \neq t$,

$$u(C_t(s)) - v\left(\frac{Y_t(s)}{w_t}\right) \geq u(C_{t'}(s)) - v\left(\frac{Y_{t'}(s)}{w_t}\right) . \tag{7}$$

The budget constraints now read as $\forall s \in S$,

$$Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq kQ(s) . \tag{8}$$

The set of implementable allocation rules is represented in the remainder of the paper by the collections $\{Q(s), Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s\in S}$, which satisfy the *C-RT-C* property, as well as the inequalities in (7) and (8). The optimal utilitarian income tax mechanism is now defined as follows.

Definition 4 With private information on taste parameters, the *optimal C-IC* income tax solves the problem of choosing $\{Q(s)$, $Y_1(s)$, $C_1(s)$, $Y_2(s)$, $C_2(s)\}_{s\in S}$, subject to the *C-RT-C* constraint, the *I-RP* constraints in (7) and the feasibility constraints in (8), in order to maximize EW.

This optimization problem differs from the one analyzed in the previous section by the presence of the *C-RT-C* constraints. Under pure assignment uncertainty, there is no need to take collective incentives into account.

5 Optimality under Collective Incentives

In this section, the properties of an optimal *C-IC* income tax are derived for each scenario. This is achieved via a two step procedure. The *first* step solves for an optimal *C-IC* income tax, taking the provision rule for the public good as given. The *second* step determines the optimal provision rule. This approach is tractable because of the fact that the *C-RT-C* constraints limit the number of admissible provision rules.

Lemma 4 Under *C-RT-C*, provision rules are increasing in both arguments, $\forall x \in \{L, H\} : Q(s_{xL}) \leq Q(s_{xH})$ and $\forall y \in \{L, H\} : Q(s_{Ly}) \leq Q(s_{Hy})$.

The monotonicity constraints stated in the lemma imply that there are only six candidate provision rules.[14] The provision rule $Q^i : Q = 0 \iff s = s_{LL}$, which is part of the informed optimum, satisfies these constraints. The same is true for provision rule $Q^{i'}$, defined by $Q = 1 \iff s = s_{HH}$, provision rule Q^1, which calls for public good provision if and only if class 1 individuals have a high taste parameter $Q^1 : Q = 1 \iff s \in \{s_{HL}, s_{HH}\}$, and the analogously defined provision rule $Q^2 : Q = 1 \iff s \in \{s_{LH}, s_{HH}\}$. Finally, the monotonicity constraints are trivially satisfied by the constant provision rules $Q \equiv 0$ and $Q \equiv 1$.
One of these six candidate provision rules is taken as given when undertaking the *first* step. The subsequent analysis focuses on the problem of finding an optimal *C-IC* income tax that implements the informed planner's provision rule Q^i. Formally, this problem is denoted *Problem* P^i and defined as follows.

The optimal *C-IC* income tax under Q^i: *Problem* P^i

An optimal *C-IC* income tax which implements provision rule Q^i solves the problem of choosing $\{Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s \in S}$ in order to maximize the expected welfare contribution from consumption and income requirements

$$\begin{aligned} EW_V := \quad & p_{LL}[V_1(s_{LL}) + V_2(s_{LL})] + p_{LH}[V_1(s_{LH}) + V_2(s_{LH})] \\ & + p_{HL}[V_1(s_{HL}) + V_2(s_{HL})] + p_{HH}[V_1(s_{HH}) + V_2(s_{HH})] \end{aligned}$$

subject to the *C-RT-C* constraints,[15]

$$\begin{aligned} & V_1(s_{LH}) = V_1(s_{HH}) \ , \quad \theta_H \geq V_1(s_{LL}) - V_1(s_{HL}) \geq \theta_L \ , \\ & V_2(s_{HL}) = V_2(s_{HH}) \ , \quad \theta_H \geq V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L \ , \end{aligned} \tag{9}$$

[14] The lemma follows from standard arguments. See the appendix.
[15] One arrives at the inequalities in (9) by plugging Q^i into the *C-RT-C* constraints.

the *I-RP* constraints in (7) and the feasibility constraints

$$\begin{aligned} &Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq 0, \quad \text{for } s = s_{LL} \ , \\ &Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq k, \quad \text{otherwise} \ . \end{aligned} \tag{10}$$

5.1 When does collective incentive compatibility matter?

With reference to problem P^i, the *Scenarios* for which the *informed optimum* survives the introduction of collective incentive requirements are easily clarified. Recall that the informed optimum is obtained by maximizing EW_V subject to *I-RP* and feasibility, without taking *C-RT-C* into account. Obviously, the informed optimum satisfies *C-RT-C* if and only if the statements in (9) remain true as one replaces $V_t(s)$ by $R_t(0)$ if $s = s_{LL}$ and by $R_t(k)$ if $s \neq s_{LL}$. That is, the informed optimum satisfies *C-RT-C* if and only if

$$\theta_H \geq R_1(0) - R_1(k) \geq \theta_L \quad \text{and} \quad \theta_H \geq R_2(0) - R_2(k) \geq \theta_L \ . \tag{11}$$

This statement coincides with the definition of *Scenario 1*, i.e. with a parameter constellation such that, at the informed optimum, *"willingness to pay for the public good is independent of earning ability."* These observations are summarized in the following proposition.

Proposition 2 The *informed optimum* has the *C-RT-C* property if and only if *Scenario 1* holds.

The informed optimum satisfies *C-RT-C* under *Scenario 1* even though, for $s = s_{LH}$ and $s = s_{HL}$, there are conflicting interests. One class of individuals – the one with the high taste parameter – wants to have the public good, while the other class opposes provision. However, this conflict does not cause collective incentive problems. The class with a high taste parameter behaves truthfully in order to ensure provision. Likewise, the class with a low taste wants to avoid provision and hence does not deviate from the truth. Under *Scenarios 2* and *3*, at least one of these properties is violated.

5.2 How to deviate from the informed optimum?

According to Proposition 2, under *Scenarios 2* and *3* collective incentive problems force a deviation from the informed optimum. To understand the planner's assessment of conceivable deviations, a characterization of the *I-RP* constrained Pareto-frontier in a neighborhood of the *informed optimum* is needed. To this end, the following problem is considered.

$$\begin{aligned}
&\max_{C_1,Y_1,C_2,Y_2} && u(C_1) - v\left(\frac{Y_1}{w_1}\right) \\
&\text{s.t.} && Y_1 - C_1 + Y_2 - C_2 \geq r \quad \text{(BC)}\ , \\
& && u(C_1) - v\left(\frac{Y_1}{w_2}\right) \leq \bar{V}_2 \qquad \text{(I-RP}_2\text{)}\ , \\
& && u(C_2) - v\left(\frac{Y_2}{w_2}\right) = \bar{V}_2\ .
\end{aligned} \tag{12}$$

I denote by $P(\bar{V}_2, r)$ the utility level of class 1 individuals that is induced by solution to problem (12).

Lemma 5 Let $v(\cdot)$ satisfy Assumption 2.

i) For all $\bar{V}_2$ and all r, Problem (12) has a unique solution. This solution is such that (BC) is binding and there is *no distortion at the top*.

ii) For all r, P is a continuous and strictly concave function of $\bar{V}_2$ with a unique maximum. For $\bar{V}_2 = R_2(r)$ – i.e. at the *informed optimum* – P is strictly decreasing in $\bar{V}_2$.

iii) For all r, there is a maximal value $\hat{R}_2(r)$ such that for $\bar{V}_2 < \hat{R}_2(r)$, (I-RP$_2$) is binding, implying a *distortion at the bottom*. For $\bar{V}_2 > \hat{R}_2(r)$, (I-RP$_2$) is not binding, and there is *no distortion at the bottom*.

Part ii) of Lemma 5 shows that there is a well defined range of parameters such that there is indeed a tradeoff between the utility of the "rich" and the utility of the "poor".[16] Moreover the *informed utilitarian optimum* does not lie at the boundary of the region where the tradeoff prevails. That is, while the utilitarian planner expands redistribution up to a level that gives rise to incentive problems – recall that the *informed optimum* has a binding *I-RP* constraint for class 2 – she does not aim at the maximal level of incentive compatible redistribution.

5.3 Scenario 2

In the following, the optimal *C-IC* income tax for Scenario 2 is analyzed. First, *Problem* P^i is solved. Then, the circumstances under which a utilitarian planner indeed wants to stick to provision rule Q^i under *C-RT-C* constraints are clarified.

[16]This is not trivial as there is a region where both classes can be made better off if $\bar{V}_2$ is increased. In that region, the potential utility gain from the fact that less resources are needed to generate a utility level of $\bar{V}_2$ is overcompensated by the utility loss from a more severe *distortion at the bottom*. See the appendix for a mathematical formulation.

Problem P^i under Scenario 2

Under Scenario 2, *C-RT-C* of the *informed optimum* fails as the more productive want to induce public good provision even if $\theta_2 = \theta_L$, i.e. the preferences of class 2 individuals cause a violation of the *independence* condition (11), and one may thus think of class 2 as the *source* of collective incentive problems. Proposition 3 characterizes the optimal utilitarian reaction to this problem.

Proposition 3 Let $v(\cdot)$ satisfy Assumption 2. Let the parameters θ_L and θ_H be such that *Scenario 2* arises. There exists $\bar{\theta}_L$ such that if $\theta_L \leq \bar{\theta}_L$, then a solution to *Problem P^i* has the following properties:

$$V_1(s_{LL}) < R_1(0) \text{ and } V_2(s_{LL}) > R_2(0) \; ;$$

$$V_1(s_{LH}) = V_1(s_{HL}) = V_1(s_{HH}) > R_1(k) \quad \text{and}$$

$$V_2(s_{LH}) = V_2(s_{HL}) = V_2(s_{HH}) < R_2(k) \; .$$

For all s, $V_1(s) = P(V_2(s), kQ^i(s))$, and there is a *distortion at the bottom.* The *C-RT-C* constraint $V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L$ for class 2 is binding, and the *C-RT-C* constraints $\theta_H \geq V_1(s_{LL}) - V_1(s_{HL}) \geq \theta_L$ for class 1 are not binding.

Under Scenario 2, the informed optimum is not achievable, as class 2 individuals have a positive willingness to pay for the public good in any state s. As the utility loss from public good provision is not large enough, class two individuals will never admit a low taste realization. To prevent a collective deviation from truth-telling, the planner has to deviate from the *informed optimum* such that, from the perspective of the "rich" class, the utility loss from public good provision goes up. This requires an increase in the level of redistribution as compared to the *informed optimum* in states with public good provision and a reduction in the level of redistribution in states with non-provision. Hence, in state s_{LL}, in which the public good is not installed, class 2 individuals receive a *C-Y* pair that generates a utility level above $R_2(0)$. In all other states, the public good is installed and class 2 gets a *C-Y* pair that implies a utility level below $R_2(k)$. These incentive corrections are chosen such that the deviation from the *informed optimum* is as small as possible in welfare terms. Consequently, the *C-RT-C* constraint $V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L$ for class 2 is binding.

The deviations from the informed optimum proceed along the *I-RP* constrained Pareto frontier; that is, class 1 individuals are made as well off as possible, given the need to fix the collective incentive problem that stems from class 2 individuals. In particular, this implies that the less productive can be made better off relative to the *informed optimum* in states with public good provision. As class 2 individuals receive a utility level below $R_2(k)$, this leaves room to raise the utility of class 1 individuals above $R_1(k)$. Analogously, in states without public good provision, class 1 individuals are worse off. As the utility level of the "rich" class exceeds $R_2(0)$, a utility level of $R_1(0)$ is out of reach for the "poor" class.

The main reason why Proposition 3 requires θ_L not to exceed some upper bound $\bar{\theta}_L$, is the requirement that the *C-RT-C* constraints of the less productive individuals are not binding.[17] The correction of redistribution claimed by Proposition 3 implies that the utility difference $V_1(s_{LL})-V_1(s_{HL})$ shrinks relative to the *informed optimum.* If the parameter θ_L is small, then there is enough room for such an adjustment. That is, the incentive corrections required to solve P^i do not induce the less productive to prefer $Q=1$ just in order to prevent the reduction of transfers that accompanies $Q=0$.

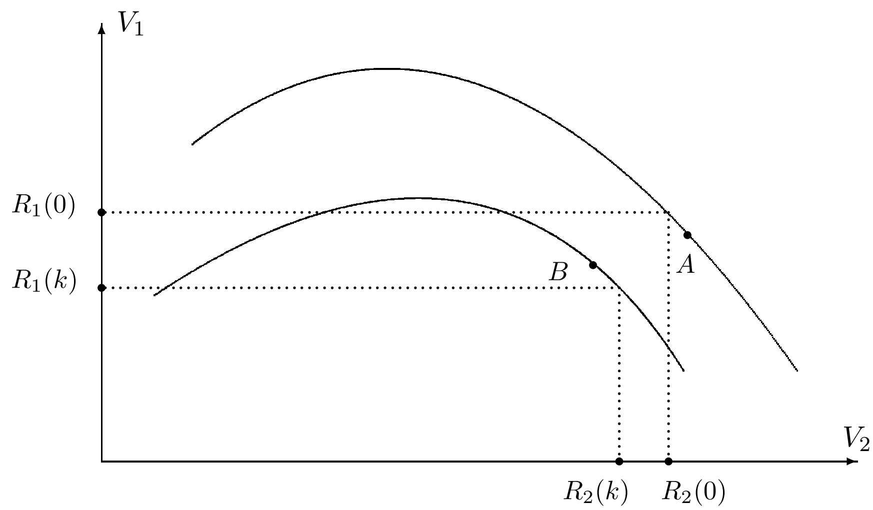

Figure 1: The graph shows the *I-RP* constrained Pareto frontiers for the revenue requirements 0 and k, respectively. Under *Sc. 2*, the difference $R_2(0)-R_2(k)$ is too small to satisfy *C-RT-C* for class 2. Under a *modest* incentive problem, the planner deviates to points A and B. Under a *severe* incentive problem, the vertical distance between these points is smaller than θ_L.

These considerations suggest the following terminology for a characterization of collective incentive problems. If $\theta_L \leq \bar{\theta}_L$, incentive problems are *modest* in the sense that it is possible to correct for the "original" collective

[17]There is also a more subtle reason. Proposition 3 claims that the *I-RP* constraints for class 2 are binding in all states. As is shown in the appendix, this is ensured if θ_L is sufficiently small. However, the logic of the proof does not rely on binding *I-RP* constraints of class 2 individuals, but on the shape of the Parteo frontier. As follows from Lemma 5, this shape is not affected as one enters the region where (I-RP$_2$) ceases to be binding.

incentive problem which stems from class 2 individuals, without creating a new one resulting from class 1 individuals. By contrast, collective incentive problems are called *severe* if a solution to P^i has two binding *C-RT-C* constraints. Here, *severity* refers to the fact that the attempt to restore *C-RT-C* for one class of individuals, renders collective manipulations attractive for the other class.

Definition 5 Denote by $\{V_1^{**}(s)\}_{s\in S}$ the utility levels realized by class 1 individuals at a solution to problem P^i. Collective incentive problems under *Scenario 2* are called *modest* if

$$V_1^{**}(s_{LL}) - V_1^{**}(s_{HL}) \quad > \quad \theta_L \ .$$

Otherwise collective incentive problems under *Scenario 2* are called *severe*.

Is Q^i the optimal provision rule under *C-RT-C* constraints?

I now turn to the question whether a utilitarian planner who faces *C-RT-C* constraints indeed wants to implement provision rule Q^i. Possibly, the welfare burden of having to adjust the transfer system if Q^i is chosen is such that a different provision rule turns out to be superior, e.g. one alternative scheme is to install the public good in every state of the world $Q(s) = 1$ for all $s \in S$. While this provision rule has the disadvantage that resources are used to cover the cost of provision even if $s = s_{LL}$, there is no need to ask individuals about their taste parameters. Hence, there is no need to deviate from the utility levels $R_1(k)$ and $R_2(k)$, which result from the *informed optimum* if the revenue requirement equals k.
In case of a *modest* incentive problem, it depends on the planner's prior whether or not provision rule Q^i is chosen. To see this, suppose first that p_{LL} is very small. Then the provision rule $Q \equiv 1$ seems attractive, as the state in which a deviation from the *informed optimum* occurs is very unlikely, i.e. the smaller p_{LL}, the more attractive provision rule $Q \equiv 1$ becomes in comparison to Q^i. As the welfare assessment EW is continuous in the prior probabilities, there must exist prior probabilities for which $Q \equiv 1$ is superior.
Now suppose that the parameters θ_L and θ_H are such that only a "small" deviation from the informed optimum is needed to achieve collective incentive compatibility – in terms of Figure 1, the points A and B are very close to the informed optimum. In such a case, the adjustments of the transfer system, required under Q^i, are negligible in welfare terms. Consequently, one may find priors such that this provision rule remains the optimal one.
In contrast, under a *severe* incentive problem, Q^i will not be chosen. To see this, suppose that the *C-RT-C* constraints

$$\theta_L \leq V_1(s_{LL}) - V_1(s_{HL}) \quad \text{and} \quad \theta_L \leq V_2(s_{LL}) - V_2(s_{LH})$$

are both binding. The two binding incentive constraints imply that all individuals are indifferent between public good provision and non-provision if $s = s_{LL}$, i.e. given $\{V_1(s), V_2(s)\}_{s \in S}$, all individuals are indifferent between the provision rules Q^i and $Q \equiv 1$. However, $Q \equiv 1$ avoids any departure from $R_1(k)$ and $R_2(k)$, implying that utilitarian welfare is higher in every state of the world. These considerations are summarized in the following proposition.

Proposition 4 Let $v(\cdot)$ satisfy Assumption 2. Let the parameters θ_L and θ_H be such that *Scenario 2* arises.

i) If collective incentive problems are *modest*, then there exist prior beliefs p such that Q^i is part of an optimal *C-IC* income tax mechanism.

ii) If collective incentive problems are *severe*, then there do not exist prior beliefs such that Q^i is part of an optimal *C-IC* income tax mechanism.

I do not discuss in more detail which of the six candidate provision rules may be supported by some prior beliefs as part of an optimal income tax mechanism. This would require for each of these candidate provision rules an analysis similar to the one conducted for Q^i; that is, one would have to determine, for each of them, the pattern of binding *C-RT-C* constraints and the welfare implications of those binding constraints.

The main results are summarized as follows: if provision rule Q^i – or any other rule that makes the decision on provision dependent on the preferences of class 2 individuals – is chosen for implementation, the planner has to accept the necessity of excessive redistribution if the public good is installed, and suboptimal redistribution if not. This may imply that the planner prefers a different provision rule in order to limit the deviations from the allocation of private goods prescribed by the *informed optimum*.

5.4 Discussion of Scenario 3

For the sake of completeness, I briefly discuss how these considerations have to be modified under *Scenario 3*. Under this parameter constellation, at the *informed optimum*, class 1 individuals oppose public good provision in any state s and class 2 individuals desire provision in any state s, i.e. there are two sources of collective incentive problems. In order to ensure collective truth-telling of class 1, at a solution to *Problem* P^i, the attractiveness of public good provision has to be increased relative to the *informed optimum*. Simultaneously for class 2, the attractiveness of public good provision has to be decreased.

Fortunately, these incentive corrections tend to complement each other. To see this, recall the properties of a solution to *Problem* P^i under *Scenario 2*,

which was dealing only with the collective incentive problem for class 2 individuals. This solution deviates from the *I-RP* constrained Pareto frontiers for revenue requirements 0 and k, respectively, such that that the utility difference between provision and non-provision shrinks for class 1 individuals relative to the *informed optimum*, i.e. this incentive correction points in the right direction as it makes public good provision more attractive from the perspective of class 1. Hence, under *Scenario 3* the solution of *Problem* P^i may be such that the *C-RT-C* constraint for class 1 is not binding. In this case, the solution of *Problem* P^i is again characterized by Proposition 3.
More generally, one has to distinguish between modest and severe incentive problems. Collective incentive problems are *modest* if, at a solution to *Problem* P^i, only the *C-RT-C* constraint for one class is binding. Otherwise they are called *severe*. These collective incentive problems may imply that Q^i is not part of an optimal *C-IC* income tax mechanism.

6 Concluding Remarks

The analysis has shown that an optimal utilitarian income tax is robust to the introduction of a *free-rider problem* on public good provision if and only if "willingness to pay for the public good" is independent of earning ability. Otherwise, collective incentive considerations force a deviation from the optimal tax scheme. Such a deviation can take different forms, a modification of the provision rule, an adjustment of the private goods allocation accompanying a given provision rule or both. The exact pattern depends on the interaction of prior probabilities and the intensity of the collective incentive problem.
This raises the question how to assess these deviations from a welfare perspective. As the analysis has shown, it is possible that those deviations make one class better off while hurting the other class, i.e. that they do not cause a departure from constrained efficiency. However, they place an additional welfare cost on redistribution if the allocation mechanism in addition has to achieve a surplus maximizing decision on public good provision. If the latter requires that, say, the "rich" admit a low valuation of public goods, then one can not simultaneously have an excessive level of redistribution in response to such a low valuation. Consequently, one has to tradeoff the utilitarian welfare gains from a more favorable solution of the *equity-efficiency tradeoff* with those from a more favorable solution of the *free-rider problem*. This tradeoff is solved such that a deviation from an *optimal income tax*, as typically defined in the literature, is desirable in order to improve the possibility to aggregate information on the willingness to pay for public goods.

A Appendix

Proof of Lemma 1: To proof the only if-part, note that, because preferences satisfy the separability property stated in equation (1), the *NDT-U*-property is an implication of *I-IC*. Obviously *I-RP* is also an implication of *I-IC*. To prove the if-part, suppose an *NDT-U* and *I-RP* allocation rule is not *I-IC*. Then there exist (w, θ) and $(\hat{w}, \hat{\theta})$ and s such that

$$u(C(w,\theta,s)) - v\left(\frac{Y(w,\theta,s)}{w}\right) < u(C(\hat{w},\hat{\theta},s)) - v\left(\frac{Y(\hat{w},\hat{\theta},s)}{w}\right) .$$

Using *NDT-U* and *I-RP* one has:

$$u(C(\hat{w},\hat{\theta},s)) - v\left(\frac{Y(\hat{w},\hat{\theta},s)}{w}\right) = u(C(\hat{w},\theta,s)) - v\left(\frac{Y(\hat{w},\theta,s)}{w}\right)$$

$$\leq u(C(w,\theta,s)) - v\left(\frac{Y(w,\theta,s)}{w}\right) .$$

Hence, a contradiction.

■

Proof of Lemma 3:
Claim 1.

$$\frac{dY_2^*(r)}{dr} > 0 \quad ; \quad \frac{dC_2^*(r)}{dr} < 0 \, .$$

Proof. These comparative statics are derived as follows: knowing that, at a solution to problem (4), the *I-RP*-constraint for type 2, as well as the budget constraint is binding allows us to setup the Lagrangean for the planner's problem. The first order conditions imply the following system of equations:

$$\frac{u'(C_1^*(r))}{u'(C_2^*(r))} = \frac{(1 - MRS_1^*) + (1 - \widehat{MRS}^*)}{MRS_1^* - \widehat{MRS}^*} , \tag{13}$$

where $\widehat{MRS}^* := \frac{1}{w_2} v'\left(\frac{Y_1^*(r)}{w_2}\right) / u'(C_1^*(r))$;

$$Y_1^*(r) - C_1^*(r) + Y_2^*(r) - C_2^*(r) = r \, , \tag{14}$$

$$u'(C_2^*(r)) = \frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right) , \tag{15}$$

$$u(C_1^*(r)) - v\left(\frac{Y_1^*(r)}{w_2}\right) = u(C_2^*(r)) - v\left(\frac{Y_2^*(r)}{w_2}\right) \ . \tag{16}$$

Differentiating these equations with respect to r yields a system of equations that can be used to solve for the derivatives of $Y_1^*(r), C_1^*(r), Y_2^*(r)$ and $C_2^*(r))$ with respect to r. After some lengthy calculations, one finds that

$$\frac{dC_2^*(r)}{dr} = \frac{(\alpha-\delta)+\epsilon(1-\delta)}{(\alpha-\delta)(\gamma-1)+\beta(1-\delta)} \ ,$$

$$\frac{dY_2^*(r)}{dr} = \gamma\frac{(\alpha-\delta)+\epsilon(1-\delta)}{(\alpha-\delta)(\gamma-1)+\beta(1-\delta)} \ ,$$

where $\alpha :=$

$$\frac{u''(C_1^*(r))\left[2u'(C_2^*(r)) + \frac{1}{w_2}v'\left(\frac{Y_1^*(r)}{w_2}\right) - \frac{1}{w_1}v'\left(\frac{Y_1^*(r)}{w_1}\right)\right]}{\frac{1}{w_1^2}v''\left(\frac{Y_1^*(r)}{w_1}\right)\left[u'(C_1^*(r)) + u'(C_2^*(r))\right] - \frac{1}{w_2^2}v''\left(\frac{Y_1^*(r)}{w_2}\right)\left[u'(C_1^*(r)) - u'(C_2^*(r))\right]}$$

and $\beta :=$

$$\frac{u''(C_2^*(r))\left[2u'(C_1^*(r)) - \frac{1}{w_2}v'\left(\frac{Y_1^*(r)}{w_2}\right) - \frac{1}{w_1}v'\left(\frac{Y_1^*(r)}{w_1}\right)\right]}{\frac{1}{w_1^2}v''\left(\frac{Y_1^*(r)}{w_1}\right)\left[u'(C_1^*(r)) + u'(C_2^*(r))\right] - \frac{1}{w_2^2}v''\left(\frac{Y_1^*(r)}{w_2}\right)\left[u'(C_1^*(r)) - u'(C_2^*(r))\right]} \ .$$

Note that, by Assumption 2, the common denominator of α and β is strictly positive. The numerator of β is negative because of the *distortion at the bottom* and the single crossing property, which imply that:

$$\frac{1}{w_2}v'\left(\frac{Y_1^*(r)}{w_2}\right) < \frac{1}{w_1}v'\left(\frac{Y_1^*(r)}{w_1}\right) < u'(C_1^*(r)) \ .$$

To see that the numerator of α is negative as well, note that equation (13) implies:

$$2u'(C_2^*(r)) + \frac{1}{w_2}v'\left(\frac{Y_1^*(r)}{w_2}\right) - \frac{1}{w_1}v'\left(\frac{Y_1^*(r)}{w_1}\right) =$$

$$\frac{\left[\frac{1}{w_1}v'\left(\frac{Y_1^*(r)}{w_1}\right)\right]^2 - \left[\frac{1}{w_2}v'\left(\frac{Y_1^*(r)}{w_2}\right)\right]^2}{2u'(C_1^*(r)) - \frac{1}{w_2}v'\left(\frac{Y_1^*(r)}{w_2}\right) - \frac{1}{w_1}v'\left(\frac{Y_1^*(r)}{w_1}\right)} > 0 \ .$$

Further,

$$\gamma := \frac{u''(C_2^*(r))}{\frac{1}{w_2^2}v''\left(\frac{Y_2^*(r)}{w_2}\right)} \leq 0 \ ,$$

$$\delta \;:=\; \frac{u'(C_1^*(r)) + \frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right)}{\frac{1}{w_2} v'\left(\frac{Y_1^*(r)}{w_2}\right) + \frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right)} \;\geq\; 1\,,$$

$$\epsilon \;:=\; \frac{\frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right)}{\frac{1}{w_2} v'\left(\frac{Y_1^*(r)}{w_2}\right) + \frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right)} \;\in\;]0,1[\,.$$

The stated properties of α, β, γ, δ and ϵ imply that Claim 1 holds true.
Claim 2.

$$0 \;>\; \frac{d}{dr}\left[u(C_2^*(r)) - v\left(\frac{Y_2^*(r)}{w_2}\right)\right] \;>\; \frac{d}{dr}\left[u(C_1^*(r)) - v\left(\frac{Y_1^*(r)}{w_1}\right)\right]\,.$$

Proof. Recall that at a solution of problem (4), the *I-RP*-constraint of the more productive type is binding (see equation (16)). Hence, it must be the case that:

$$\frac{d}{dr}\left[u(C_2^*(r)) - v\left(\frac{Y_2^*(r)}{w_2}\right)\right] \;=\; \frac{d}{dr}\left[u(C_1^*(r)) - v\left(\frac{Y_1^*(r)}{w_2}\right)\right]\,.$$

Due to the convexity of $v(\cdot)$ and $\frac{dY_1^*(r)}{dr} > 0$, one also has:

$$\frac{d}{dr}\left[u(C_1^*(r)) - v\left(\frac{Y_1^*(r)}{w_2}\right)\right] \;>\; \frac{d}{dr}\left[u(C_1^*(k)) - v\left(\frac{Y_1^*(r)}{w_1}\right)\right]\,.$$

To see that also the first inequality holds, note that, using (15), one has:

$$\frac{d}{dk}\left[u(C_2^*(r)) - v\left(\frac{Y_2^*(r)}{w_2}\right)\right] \;=\; u'(C_2^*(r))\left[\frac{dC_2^*(r)}{dr} - \frac{dY_2^*(r)}{dr}\right]\,.$$

This expression is strictly negative by *Claim 1.*

■

Proof of Proposition 1: Before proceeding with the proof, an additional piece of notation is introduced. Recall that an income tax mechanism specifies for each s the variables $Q(s)$ and $C_1(w_1, \theta_L, s)$, $Y_1(w_1, \theta_L, s)$, $C_1(w_1, \theta_H, s)$, $Y_1(w_1, \theta_H, s)$ and $C_2(w_2, \theta_L, s)$, $Y_2(w_2, \theta_L, s)$, $C_2(w_2, \theta_H, s)$, $Y_2(w_2, \theta_H, s)$. I henceforth refer to this list of variables, as the allocation $A(s)$ for state s. An income tax mechanism M can hence be summarized as a list $M = (A(s_{LL}),\ A(s_{LH}),\ A(s_{HL}),\ A(s_{HH}))$. If I want to describe an income tax mechanism M' that, say, coincides with a *predefined* income tax mechanism M in all states except s_{LL} and chooses in this state the allocation prescribed by M for state s_{LH}, I write $M' = (A(s_{LH}),\ A(s_{LH}),\ A(s_{HL}),\ A(s_{HH}))$.

Claim 1. Consider an income tax mechanism M. Suppose there are no pooling outcomes. Then, there is no manipulating coalition that misreports productivity parameters.

Proof. It is first shown that there is no coalition that contains individuals of both classes who both misreport productivity. Suppose, to the contrary, that there exist $s \in S$, $J \in \mathcal{J}$, containing individuals of both types and $\delta(\hat{\gamma}_J, s) \in \mathcal{D}$ such that individuals of both classes misreport productivity and such that $\forall j \in J$:

$$\theta^j Q(\hat{s}) + u(C(w^j, \theta^j, \hat{s})) - v\left(\frac{Y(w^j, \theta^j, \hat{s})}{w^j}\right)$$

$$= \theta^j Q(\hat{s}) + u(C(\hat{w}^j, \hat{\theta}^j, \hat{s})) - v\left(\frac{Y(\hat{w}^j, \hat{\theta}^j, \hat{s})}{w^j}\right)$$

$$> \theta^j Q(s) + u(C(w^j, \theta^j, s)) - v\left(\frac{Y(w^j, \theta^j, s)}{w^j}\right) .$$

Evaluating this condition for both types and using the *NDT-U* property implies that for all $t \in \{1, 2\}$,

$$u(C(w_1, \theta_L, \hat{s})) - v\left(\frac{Y(w_1, \theta_L, \hat{s})}{w_t}\right) = u(C(w_1, \theta_H, \hat{s})) - v\left(\frac{Y(w_1, \theta_H, \hat{s})}{w_t}\right)$$

$$= u(C(w_2, \theta_L, \hat{s})) - v\left(\frac{Y(w_2, \theta_L, \hat{s})}{w_t}\right) = u(C(w_2, \theta_H, \hat{s})) - v\left(\frac{Y(w_2, \theta_H, \hat{s})}{w_t}\right) .$$

Due to the single crossing property, those equalities hold for all t only if

$$(C(w_1, \theta_L, \hat{s}), Y(w_1, \theta_L, \hat{s})) = (C(w_1, \theta_H, \hat{s}), Y(w_1, \theta_H, \hat{s}))$$

$$= (C(w_2, \theta_L, \hat{s}), Y(w_2, \theta_L, \hat{s})) = (C(w_2, \theta_H, \hat{s}), Y(w_2, \theta_H, \hat{s})) .$$

Hence, this contradicts the assumption that there is no pooling. We may thus assume that all individuals of one class reveal their productivity parameter. But then non-detectability requires that all individuals of the other class reveal their productivity parameter as well. Otherwise the announced distribution would not be compatible with the commonly known fact that half of the population has earning ability $w_{t'}$ and half of the population has earning ability w_t.

Claim 2. Consider an income tax mechanism M. Suppose there are no pooling outcomes. Suppose the induced utility allocation $\{\tilde{U}_1(s), \tilde{U}_2(s)\}_{s \in S}$ has the *C-RT-C* property. Then, there is no manipulating coalition that contains individuals of only one type.

Proof. Suppose all individuals of type t' reveal their characteristics truthfully. By *Claim 1*, all individuals of type t, $t \neq t'$ reveal their earning ability truthfully. Moreover, all individuals of type t have to report the same taste parameter. Otherwise the announced distribution does not belong to $\mathcal{D}$. Hence, any conceivable manipulation must involve a revelation of earning ability and a collective misreport of taste on the class level. Those manipulations are ruled, by the *C-RT-C* property.

Claim 3. Consider an income tax mechanism $M = (A(s_{LL}), A(s_{LH}), A(s_{HL}), A(s_{HH}))$. Suppose there are no pooling outcomes. Suppose the induced utility allocation $\{\tilde{U}_1(s), \tilde{U}_2(s)\}_{s \in S}$ has the *C-RT-C* property and is Pareto-optimal within the set of utility allocations which are implementable by an income tax mechanism with the *C-RT-C* property. Then, there is no manipulating coalition that contains individuals of both types.
Proof. By *Claim 1*, individuals of both types reveal their productivity parameters. By non-detectability, all individuals who announce the same productivity parameter also have to announce the same taste parameter. Hence, the only conceivable manipulation that contains individuals of both types is such that all type 1 individuals misreport θ_1 and all type 2 individuals misreport θ_2. I show in the following, that Pareto-optimality within the set of *C-RT-C* utility allocations implies that there does not exist a stable joint manipulation of taste parameters that makes both type 1 and type 2 individuals strictly better off.
The proof proceeds by contradiction. Suppose there is a a joint lie on taste parameters that makes both type 1 and type 2 individuals strictly better off. Without loss of generality, assume that the true state is s_{LL}.[18] An undetectable joint collective lie induces the state perception s_{HH}. Such a collective lie makes all coalition members better off only if

$$\begin{aligned} \theta_L Q(s_{LL}) + V_1(s_{LL}) &< \theta_L Q(s_{HH}) + V_1(s_{HH}) \, , \\ \theta_L Q(s_{LL}) + V_2(s_{LL}) &< \theta_L Q(s_{HH}) + V_2(s_{HH}) \, . \end{aligned} \tag{17}$$

Due to *NDT-U*, this collective deviation is individually stable, as it does only involve misreports of taste parameters. To achieve collective stability as well, the following *C-RT-C* constraints have to be binding,

$$\begin{aligned} \theta_L Q(s_{LH}) + V_1(s_{LH}) &= \theta_L Q(s_{HH}) + V_1(s_{HH}) \, , \\ \theta_L Q(s_{HL}) + V_2(s_{HL}) &= \theta_L Q(s_{HH}) + V_2(s_{HH}) \, ; \end{aligned} \tag{18}$$

otherwise a coalition consisting only of type 1 individuals or a coalition

[18] The reasoning that follows is applicable for any conceivable constellation under which type 1 and type 2 individuals might consider a joint manipulation of taste parameters.

consisting only of type 2 individuals would want to deviate once more after the state perception s_{HH} has been induced.

(a) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &\geq \theta_H Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &\geq \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{19}$$

By (17), the following income tax mechanism $M' = (A(s_{HH}), A(s_{LH}), A(s_{HL}), A(s_{HH}))$ is Pareto superior. It is easily verified that M' satisfies all *C-RT-C* constraints if (19) holds.

(b) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &< \theta_H Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &< \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{20}$$

By (17), (18) and (20), the following income tax mechanism $M' = (A(s_{HH}), A(s_{HH}), A(s_{HH}), A(s_{HH}))$ is Pareto superior. Obviously, M' satisfies all *C-RT-C* constraints.

(c) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &< \theta_H Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &\geq \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{21}$$

The following income tax mechanism $M' = (A(s_{HH}), A(s_{LH}), A(s_{HH}), A(s_{HH}))$ is Pareto superior. It follows from (17) that both types are better off in state s_{LL}. It follows from (18) that type 2 is not worse off in state s_{HL}, and it follows from (21) that type 1 is strictly better off in state s_{HL}. Moreover, it is easily verified that M' satisfies all *C-RT-C* constraints.

(d) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &\geq \theta_H Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &< \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{22}$$

Then, along the same lines as under (c), one shows that $M' = (A(s_{HH}), A(s_{HH}), A(s_{HL}), A(s_{HH}))$ is Pareto superior and satisfies *C-RT-C*.

Claims 1 to 3 are summarized as follows: Suppose there are no pooling outcomes. Then a utility allocation that is Pareto-optimal within the set of utility allocations which are implementable by an income tax mechanism with the *C-RT-C* property also possesses the (more demanding) *C-IC* property. We have thus shown that the set of Pareto-optimal utility allocations under *C-IC* income tax mechanisms coincides with the set of Pareto-optimal utility allocations under *C-RT-C* income tax mechanisms. In the following, I may hence focus on the latter set. Recalling Lemma 1, these utility allocations are achievable by means of a feasible allocation mechanism with the *I-RP*, the *NDT-U* and the *C-RT-C* property. The following claim remains to be established:

Claim 4. If pooling can be excluded, a Pareto-optimal utility allocation is implementable if and only if it is implementable by a feasible allocation mechanism which satisfies the *NDT-CY*, the *I-RP* and the *C-RT-C* property.

Proof. As the *NDT-CY* property implies the *NDT-U* property, the if-part is trivial. To prove the only if-part, consider an implementable utility allocation and the underlying coalition-proof income tax mechanism $Q(\cdot), C(\cdot), Y(\cdot)$. Construct an allocation rule $Q'(\cdot), C'(\cdot), Y'(\cdot)$ which has the *NDT-CY* property and coincides with $Q(\cdot), C(\cdot), Y(\cdot)$ "on the equilibrium path" as follows: $\forall s \in S$, $Q'(s) = Q(s)$ and $\forall (w_t, \theta) \in \Gamma_r(s)$, $C'(w_t, \theta, s) = C(w_t, \theta, s)$ and $Y'(w_t, \theta, s) = Y(w_t, \theta, s)$. For $\theta' \neq \theta$, $C'(w_t, \theta', s) = C'(w_t, \theta, s)$ and $Y'(w_t, \theta', s) = Y'(w_t, \theta, s)$. By construction, $Q'(\cdot), C'(\cdot), Y'(\cdot)$ is feasible and inherits the *NDT-U*, the *I-RP* and the *C-RT-C*-property from $Q(\cdot), C(\cdot), Y(\cdot)$.

■

Proof of Lemma 4: Consider for example the *C-RT-C* constraints for class 1 given that $\theta_2 = \theta_L$:

$$\text{if } \theta_1 = \theta_L : \quad \theta_L Q(s_{LL}) + V_1(s_{LL}) \geq \theta_L Q(s_{HL}) + V_1(s_{HL}) \ ,$$

$$\text{if } \theta_1 = \theta_H : \quad \theta_H Q(s_{HL}) + V_1(s_{HL}) \geq \theta_H Q(s_{LL}) + V_1(s_{LL}) \ .$$

Adding up these inequalities gives $Q(s_{HL}) \geq Q(s_{LL})$. Similarly, one derives the constraints $Q(s_{LH}) \geq Q(s_{LL})$, $Q(s_{HH}) \geq Q(s_{HL})$ and $Q(s_{HH}) \geq Q(s_{LH})$.

■

Proof of Lemma 5: The argument is only sketched. Consider the Lagrangean of problem (12):

$$\begin{aligned}\mathcal{L} &= u(C_1) - v\left(\frac{Y_1}{w_1}\right) - \mu[k + C_1 + C_2 - Y_1 - Y_2] \\ &\quad -\lambda[u(C_1) - v\left(\frac{Y_1}{w_2}\right) - \bar{V}_2] - \nu[\bar{V}_2 - u(C_2) - v\left(\frac{Y_1}{w_2}\right)] \, .\end{aligned}$$

Deriving first order conditions, one easily verifies that *(BC)* has to be binding, that there is *no distortion at the top* and a *distortion at the bottom* if and and only if (I-RP$_2$) is binding.
Denote the solution of problem (12) by $(\bar{Y}_1(\bar{V}_2), \bar{C}_1(\bar{V}_2), \bar{Y}_2(\bar{V}_2), \bar{C}_2(\bar{V}_2))$. Uniqueness of this solution can be established as follows. Strict quasiconcavity of preferences and the property of *no distortion at the top* uniquely determine $\bar{Y}_2$ and $\bar{C}_2$ as a function of $\bar{V}_2$. The fact that *(BC)* is binding yields a unique iso-tax-revenue line $Y_1 - C_1 = \gamma$, with $\gamma = C_2(\bar{V}_2) - Y_2(\bar{V}_2) + r$, on which the point $(\bar{Y}_1(\bar{V}_2), \bar{C}_1(\bar{V}_2))$ can be found. More precisely, $(\bar{Y}_1(\bar{V}_2), \bar{C}_1(\bar{V}_2))$ maximizes $u(C_1) - v(Y_1/w_1)$ subject to I-RP$_2$ and $Y_1 - C_1 = \gamma$. Again, due to strict quasiconcavity, the latter problem has a unique solution.
Denote the optimal values of the multipliers at the solution of problem (12) by $\bar{\lambda}(\bar{V}_2)$ and $\bar{\nu}(\bar{V}_2)$. These multipliers are used to study how P depends on $\bar{V}_2$. The following property is used:[19]

$$\frac{\partial P}{\partial \bar{V}_2} = \bar{\lambda}(\bar{V}_2) - \bar{\nu}(\bar{V}_2) \, .$$

Similarly as for the proof of Lemma 3, comparative statics of the solution of problem (12) with respect to $\bar{V}_2$ can be derived. Based on this exercise, the comparative statics of the Lagrangean multipliers can be determined.[20] The details of the computations are omitted. One arrives at the following results:

(a) Suppose first that (I-RP$_2$) is binding.[21] Using Assumption 2, one verifies that the function $\bar{\lambda}(\bar{V}_2)$ decreases in $\bar{V}_2$ and that the function

[19] $\bar{\lambda}(\bar{R}_2) \geq 0$ captures the effect that a lower level of $\bar{V}_2$ tends to reduce P due to a worsening of incentive problems. The expression $-\bar{\nu}(\bar{V}_2) \leq 0$ shows that a lower level of $\bar{V}_2$ allows us to increase P as less resources are needed to equip type 2 individuals with a utility level of $\bar{V}_2$.

[20] The first order conditions imply:

$$\bar{\lambda}(\bar{V}_2) = \frac{1 - \overline{MRS}_1}{1 - \widehat{MRS}} \quad \text{and} \quad \bar{\nu}(\bar{V}_2) = \frac{u'(\bar{C}_1)}{u'(\bar{C}_2)} \frac{\overline{MRS}_1 - \widehat{MRS}}{1 - \widehat{MRS}} \, , \tag{23}$$

where $\overline{MRS}_1 := \frac{1}{w_1} v'\left(\frac{\bar{Y}_1}{w_1}\right) / u'(\bar{C}_1)$ and $\widehat{MRS} := \frac{1}{w_2} v'\left(\frac{\bar{Y}_1}{w_2}\right) / u'(\bar{C}_1)$.

[21] The existence of a value $\bar{V}_2$ such that (I-RP$_2$) is binding follows from Lemma 2.

$\bar{\nu}(\bar{V}_2)$ increases in $\bar{V}_2$, i.e. as long as (I-RP$_2$) is binding the function P is strictly concave in $\bar{V}_2$ and one has:

$$\frac{\partial^2 P}{\partial (\bar{V}_2)^2} \quad = \quad \bar{\lambda}'(\bar{V}_2) - \bar{\nu}'(\bar{V}_2) \quad < \quad 0 \, .$$

(b) Assume that (I-RP$_2$) is not binding.[22] The first order conditions imply $\bar{\lambda}(\bar{V}_2) = 0$ and $\bar{\nu}(\bar{V}_2) = u'(\bar{C}_1)/u'(\bar{C}_2)$. Again, the comparative statics with respect to $\bar{V}_2$ reveal

$$\frac{\partial^2 P}{\partial (\bar{V}_2)^2} \quad = \quad -\bar{\nu}'(\bar{V}_2) \quad < \quad 0 \, .$$

(c) As $\bar{\lambda}(\bar{V}_2)$ decreases in $\bar{V}_2$, there is a critical value $\hat{R}_2$, such that if this critical value is exceeded, (I-RP$_2$) is not binding anymore. Moreover, one can show that the function $\bar{\lambda}(\bar{V}_2) - \bar{\nu}(\bar{V}_2)$ is continuous at $\hat{R}_2$.

(d) P has a maximum as follows from the existence of a solution of the following problem:

$$\begin{aligned} \max_{C_1,Y_1,C_2,Y_2} \quad & u(C_1) - v\left(\frac{Y_1}{w_1}\right) \\ \text{s.t.} \quad & Y_1 - C_1 + Y_2 - C_2 \geq k \, , \\ & u(C_2) - v\left(\frac{Y_2}{w_2}\right) \geq u(C_1) - v\left(\frac{Y_1}{w_2}\right) \, . \end{aligned} \tag{24}$$

Denote by $\tilde{V}_2$ the utility level that results for type 2 individuals at a solution to problem (24). Using the first order conditions of problem (24) allows us to verify that $\bar{\lambda}(\tilde{V}_2) - \bar{\nu}(\tilde{V}_2) = 0$.

(e) Finally, use the first order condition (13) of the *informed problem* to substitute for $u'(\bar{C}_1)/u'(\bar{C}_2)$ in the formula for $\bar{\nu}(\bar{V}_2)$ (see (23)), one gets $\bar{\lambda}(R_2(r)) - \bar{\nu}(R_2(r)) = -1$.

■

Proof of Proposition 3: I consider a relaxed version of *Problem* P^i, referred to as *Problem* P^i_x. P^i_x takes only a subset of the constraints of P^i into account. I show below that a solution to P^i_x satisfies these neglected constraints.
Formally P^i_x is defined as follows. Maximize EW_V subject to the feasibility

[22] The existence of a value of $\bar{V}_2$ such that (I-RP$_2$) is not binding can e.g. be established by the *laissez faire* solution, where individuals of type t choose (Y_t, C_t) to maximize utility under the constraint $Y_t = C_t + \frac{k}{2}$.

constraints in (10), the *I-RP* constraints for type 2 and the following subset of the *C-RT-C* constraints:

$$V_1(s_{LH}) = V_1(s_{HH})\ ,$$

$$V_2(s_{HL}) = V_2(s_{HH})\ , \quad V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L\ .$$

$\{V^{**}_{t,x}(s)\}_{s\in S}$ denotes the utility levels realized by class t at a solution to P^i_x. The following assumption formalizes the statement in Proposition 3 that θ_L must not exceed some upper bound $\bar{\theta}_L$.

Assumption 3 $\theta_L < \min\{\hat{R}_2(0) - R_2(k), P(R_2(k) + \theta_L, 0) - P(R_2(0) - \theta_L, r)\}$.

As will become clear, $\theta_L \leq \hat{R}_2(0) - R_2(k)$ ensures that, in every state s, there is a *distortion at the bottom* at a solution of P^i_x. $\theta_L \leq P(R_2(k) + \theta_L, 0) - P(R_2(0) - \theta_L, r)$ ensures that a solution of P^i_x does not violate the neglected *C-RT-C* constraint for class 1.[23]

Under Assumption 3, Proposition 3 follows from the following observations.

(a) In every state s, the budget constraint is binding and there is *no distortion at the top.*
Proof. This follows from setting up the Lagrangean of P^i_x and deriving first order conditions.

(b) $V^{**}_{2,x}(s_{LL}) \leq \theta_L + R_2(k)$.
Proof. Suppose, to the contrary, that $V^{**}_{2,x}(s_{LL}) > \theta_L + R_2(k)$. Then the planner could choose, instead, the allocation $(Y^*_1(k), C^*_1(k), Y^*_2(k), C^*_2(k))$ for $s \neq s_{LL}$. For $s = s_{LL}$, the planner could choose $V^{**}_{2,x}(s_{LL}) = \theta_L + R_2(k)$ and $V_1 = P(\theta_L + R_2(k), 0)$. Due to the monotonicity properties established in Lemma 5, this would increase utilitarian welfare in every state s.

(c) $V^{**}_{2,x}(s_{LH}) \leq R_2(k)$.
Proof. The *C-RT-C* constraints imposed under P^i_x imply $V^{**}_{2,x}(s_{LH}) \leq V^{**}_{2,x}(s_{LL}) - \theta_L$. Combing this with (b) yields (c).

(d) $V^{**}_{1,x}(s_{LL}) = P(V^{**}_{2,x}(s_{LL}), 0)$ and $V^{**}_{1,x}(s_{HL}) = P(V^{**}_{2,x}(s_{HL}), k)$. Moreover, there is a *distortion at the bottom* in state s_{LL}.

[23] Recall that under *Scenario 2*, $\theta_L > R_2(0) - R_2(k)$. If θ_L does not exceed $R_2(0) - R_2(k)$ by too much, i.e. $\theta_L \simeq R_2(0) - R_2(k)$, then, the assumption $\theta_L < P(R_2(k) + \theta_L, 0) - P(R_2(0) - \theta_L, r)$ is satisfied. To see this: if $\theta_L \simeq R_2(0) - R_2(k)$, the continuity property established in Lemma 5 implies that $P(R_2(k) + \theta_L, 0) - P(R_2(0) - \theta_L, r) \simeq R_1(0) - R_1(k)$. By definition of *Scenario 2*, the latter term exceeds θ_L.

Proof. $V^{**}_{1,x}(s_{LL}) \neq P(V^{**}_{2,x}(s_{LL}), 0)$ or $V^{**}_{1,x}(s_{HL}) \neq P(V^{**}_{2,x}(s_{HL}), k)$ immediately yields a contradiction to optimality. The *distortion at the bottom* in state s_{LL} follows from Assumption 3 and observation (b), which imply that $V^{**}_{2,x}(s_{LL}) < \hat{R}_2(0)$.

(e) $V^{**}_{2,x}(s_{LL}) \geq R_2(0)$ and $V^{**}_{1,x}(s_{LL}) \leq R_1(0)$.
Proof. Given (d), if (e) is false, then $V^{**}_{2,x}(s_{LL}) < R_2(0)$ and $V^{**}_{1,x}(s_{LL}) > R_1(0)$. Then, using the monotonicity properties established in Lemma 5, it is possible to increase $V_2(s_{LL})$ and to decrease $V_1(s_{LL})$ along the (I-RP$_2$) constrained Pareto frontier without violating the constraint $V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L$, thereby increasing utilitarian welfare in state s_{LL}.

(f) $V^{**}_{2,x}(s_{LH}) = V^{**}_{2,x}(s_{HH}) = V^{**}_{2,x}(s_{HL}) =: V^{**}_{2,x}(s_H) \leq R_2(k)$.
Proof. $V^{**}_{2,x}(s_{HH}) = V^{**}_{2,x}(s_{HL})$ is a *C-RT-C* constraint, and $V^{**}_{2,x}(s_{LH}) \leq R_2(k)$ has been established in (c). Hence it remains to be shown that $V^{**}_{2,x}(s_{LH}) = V^{**}_{2,x}(s_{HH})$. To the contrary, let $V^{**}_{2,x}(s_{LH}) \neq V^{**}_{2,x}(s_{HH})$. Optimality requires that $V^{**}_{1,x}(s_{LH}) = V^{**}_{1,x}(s_{HH})$ is the utility level realized at a solution to the following problem:
Choose $(C_1(s_{LH}), Y_1(s_{LH}), C_2(s_{LH}), Y_2(s_{LH}))$ and $(C_1(s_{HH}), Y_1(s_{HH}), C_2(s_{HH}), Y_2(s_{HH}))$ in order to maximize

$$u(C_1(s_{LH})) - v\left(\frac{Y_1(s_{LH})}{w_1}\right)$$

subject to the following constraints: Feasibility,

$$Y_1(s_{LH}) - C_1(s_{LH}) + Y_2(s_{LH}) - C_2(s_{LH}) \geq r \qquad (\text{BC}(s_{LH}))\ ,$$

$$Y_1(s_{HH}) - C_1(s_{HH}) + Y_2(s_{HH}) - C_2(s_{HH}) \geq r \qquad (\text{BC}(s_{HH}))\ ,$$

the *I-RP* constraints for type 2,

$$u(C_1(s_{LH})) - v\left(\frac{Y_1(s_{LH})}{w_2}\right) \leq V^{**}_{2,x}(s_{LH}) \qquad (\text{I-RP}_2(s_{LH}))\ ,$$

$$u(C_1(s_{HH})) - v\left(\frac{Y_1(s_{HH})}{w_2}\right) \leq V^{**}_{2,x}(s_{HH}) \qquad (\text{I-RP}_2(s_{HH}))\ ,$$

the *C-RT-C* constraint,

$$u(C_1(s_{LH})) - v\left(\frac{Y_1(s_{LH})}{w_1}\right) = u(C_1(s_{HH})) - v\left(\frac{Y_1(s_{HH})}{w_1}\right)\ ,$$

and the requirement to deliver the following utility levels to class 2,

$$u(C_2(s_{LH})) - v\left(\frac{Y_2(s_{LH})}{w_2}\right) = V^{**}_{2,x}(s_{LH})\ ,$$

$$u(C_2(s_{HH})) - v\left(\frac{Y_2(s_{HH})}{w_2}\right) = V^{**}_{2,x}(s_{HH})\ .$$

Suppose, without loss of generality, that $V^{**}_{2,x}(s_{LH}) < V^{**}_{2,x}(s_{HH})$. One can show that a solution to this problem has the following properties:[24] $(C_1(s_{LH}), Y_1(s_{LH})) = (C_1(s_{HH}), Y_1(s_{HH}))$, $(\text{BC}(s_{HH}))$ and $(\text{I-RP}_2(s_{LH}))$ are binding, while $(\text{BC}(s_{LH}))$ and $(\text{I-RP}_2(s_{HH}))$ hold with a strict inequality. However, a strict inequality in $(\text{BC}(s_{LH}))$ contradicts (a).

(g) $V^{**}_{1,x}(s_{LH}) = V^{**}_{1,x}(s_{HH}) = V^{**}_{1,x}(s_{HL}) =: V^{**}_{1,x}(s_H) = P(V^{**}_{2,x}(s_H), k) \geq R_1(k)$, implying a *distortion at the bottom* in states s_{LH}, s_{HL} and s_{HH}. Moreover, at a solution to *Problem* P^i_x $(C_1(s_{HL}),\ Y_1(s_{HL})) = (C_1(s_{LH}),\ Y_1(s_{LH})) = (C_1(s_{HH}),\ Y_1(s_{HH})) =: (C_1(s_H),\ Y_1(s_H))$ and $(C_2(s_{HL}),\ Y_1(s_{HL})) = (C_2(s_{LH}),\ Y_1(s_{LH})) = (C_2(s_{HH}),\ Y_1(s_{HH})) =: (C_2(s_H),\ Y_2(s_H))$.
Proof. The first statement follows from (d), (f), optimality considerations and the monotonicity properties established in Lemma 5, making use of the fact that $V^{**}_{2,x}(s_H) \leq R_2(k)$. The equality of (C_t, Y_t) pairs across states follows from the uniqueness established in Lemma 5.

(h) $V^{**}_{1,x}(s_H) \neq R_1(k)$ and $V^{**}_{2,x}(s_H) \neq R_2(k)$ and $V^{**}_{2,x}(s_{LL}) \neq R_2(0)$ and $V^{**}_{1,x}(s_{LL}) \neq R_1(0)$ and $V^{**}_{2,x}(s_{LL}) - V^{**}_{2,x}(s_{LH}) = \theta_L$.
Proof. This follows from setting up the Lagrangean of P^i_x and deriving first order conditions using the above results on the pattern of binding constraints. In particular, if the constraint $V_2(s_{LL}) - V_2(s_H) \geq \theta_L$ was not binding, then the first order conditions would result in the *informed optimum*, which is known to violate this constraint. The presence of the corresponding multiplier in the first order conditions shows that, for all s, the resulting allocation differs from the one chosen by the informed planner.

(i) $\theta_H > R_1(0) - R_1(k) > V^{**}_{1,x}(s_{LL}) - V^{**}_{1,x}(s_H)$
Proof. This follows from (e) (g), (h) and the definition of *Scenario 2*.

(j) $V^{**}_{1,x}(s_{LL}) \geq P(R_2(k) + \theta_L, 0)$ and $V^{**}_{1,x}(s_H) \leq P(R_2(0) - \theta_L, r)$.
Proof. The first inequality follows from (b), (d) and the monotonicity property established in Lemma 5. The second inequality is established as follows: Analogously as in (b), one shows that $V^{**}_{2,x}(s_H) \geq R_2(0) - \theta_L$

[24] I omit the details. They involve the following steps: Show that there is *no distortion at the top* via an analysis of first order conditions. This determines $(C_2(s_{LH}), Y_2(s_{LH}))$ and $(C_2(s_{HH}), Y_2(s_{HH}))$ as functions of the utility levels $V^{**}_{2,x}(s_{LH})$ and $V^{**}_{2,x}(s_{HH})$, respectively. Secondly, show that this implies that the feasible set for a choice of $(C_1(s_{LH}), Y_1(s_{LH}))$ and $(C_1(S_{HH}), Y_1(s_{HH}))$ is effectively restricted only by $(\text{BC}(s_{HH}))$ and $(\text{I-RP}_2(s_{LH}))$. Thirdly, use the geometry of this set, the strict quasiconcavity of preferences, as well as the fact that $V^{**}_{2,x}(s_{LH}) \leq R_2(k)$ established in (c), to show that there is a unique optimal choice for both $(C_1(s_{LH}), Y_1(s_{LH}))$ and $(C_1(S_{HH}), Y_1(s_{HH}))$ and that at this solution $(\text{BC}(s_{HH}))$ and $(\text{I-RP}_2(s_{LH}))$ are binding while $(\text{BC}(s_{LH}))$ and $(\text{I-RP}_2(s_{HH}))$ are slack.

and then uses $V_{2,x}^{**}(s_H) = P(V_{2,x}^{**}(s_H), k)$ and again the monotonicity property.

(k) At solution to P_x^i, the neglected *C-RT-C* constraint for class 1 is satisfied. I.e. $\theta_H > V_{1,x}^{**}(s_{LL}) - V_{1,x}^{**}(s_H) > \theta_L$
Proof. This follows from (i), (j) and Assumption (3).

■

References

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

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

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

Bernheim, B. and Whinston, M. (1986). Coalition-proof Nash equilibria II. Applications. *Journal of Economic Theory*, 42:13–29.

Boadway, R., Cuff, K., and Marchand, M. (2000). Optimal income taxation with quasi-linear preferences revisited. *Journal of Public Economic Theory*, 2:435–460.

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

Chung, K. and Ely, J. (2004). Foundations of dominant strategy mechanisms. Mimeo, Northwestern University.

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

Farrell, J. and Rabin, M. (1996). Cheap talk. *Journal of Economic Perspectives*, 10:103–118.

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

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

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

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

Hellwig, M. (2005). A contribution to the theory of optimal income taxation. Preprint 2005/23, Max Planck Institute for Research on Collective Goods, Bonn.

Judd, K. (1985). The law of large numbers with a continuum of iid random variables. *Journal of Economic Theory*, 35:19–25.

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

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

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

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

Mirrlees, J. (1975). Optimal commodity taxation in a two-class economy. *Journal of Public Economics*, 4:27–33.

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

Stiglitz, J. (1982). Self-selection and Pareto-efficient taxation. *Journal of Public Economics*, 17:213–240.

Stiglitz, J. (1987). Pareto efficient and optimal taxation and the new new welfare economics. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*. North-Holland, Amsterdam.

Weymark, J. (1986). A reduced-form optimal nonlinear income tax problem. *Journal of Public Economics*, 30:199–217.