Optimal Income Taxation and Public Good Provision with Endogenous Interest Groups*

Felix Bierbrauer†

MPI for Research on Collective Goods, Bonn

July 20, 2007

Abstract

This paper studies public good provision when agents differ in earning abilities as well as preferences. There is aggregate uncertainty with respect to public goods preferences. Heterogeneity in skills makes redistribution desirable and generates an equity-efficiency tradeoff. A larger revenue requirement aggravates this tradeoff. Consequently, public good provision makes income transfers less desirable. High-skilled individuals thus have an incentive to exaggerate their preferences for public goods. Analogously, low-skilled individuals lobby against public good provision. A requirement of collective incentive compatibility eliminates these biases. It implies that income transfers are increased whenever a public good is provided and are decreased otherwise.

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

JEL: D71, D82, H21, H41

---

*An earlier version of this paper circulated under the title *Optimal Income Taxation and Public Good Provision in a Two-Class Economy.*

†I am very grateful for numerous discussions with and encouragement from my supervisor Martin Hellwig. I also benefited a lot from comments of Peter Diamond, Christoph Engel, Thomas Gaube, Mike Golosov, Hendrik Hakenes, Thomas Mertens, 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, the second summer school on social heterogeneity in Louvain-la-Neuve 2006, the North American Summer Meeting of the Econometric Society in Durham 2007 as well as seminar participants at the WZB Berlin and the universities of Mannheim, Cologne, Munich, Duisburg-Essen and the Massachusetts Institute of Technology.

1 Introduction

This paper studies optimal income taxation and public good provision under the assumption that individuals differ both in their earning abilities and in their valuations of a public good. Taxes are raised for two purposes. First, they are used to finance direct income transfers between high-skilled and low-skilled individuals. Second, they generate the revenues that are needed to cover the cost of public good provision. Direct income transfers and public goods are linked through the public sector budget constraint. This raises the possibility that redistribution and public good provision interfere with each other.

Many publicly provided goods have a distributional impact, in the sense that they are highly valued by some groups in the population but not by others. The quality of public schools or the extent of publicly provided health care are a major concern for individuals with low incomes. Police patrols in unsafe neighborhoods benefits those who live there, usually individuals with low income. By contrast, the quality of a country's judicial system seems to be more highly valued by those who run businesses and earn, on average, higher incomes. The use of public money for a theater, an opera building or a museum benefits those who enjoy these facilities, typically well-educated individuals with above-average incomes. Finally, public provision of higher education tends to benefit people with higher incomes.
The existing literature on optimal transfer and expenditure policies addresses these issues under the assumption that the cross-section distribution of preferences is common knowledge.[1] By contrast, the present paper focusses on information and incentive problems that arise if individuals are privately informed about their valuation of a public good. This raises a new set of issues. Suppose, for instance, that individuals want to make sure that tax revenues are used for their preferred public project. Those individuals might hence lobby for this public good in a way that makes it difficult to observe their true willingness to pay. The rivalry between the transfer system and public good provision may create its own set of incentive problems. Individuals with a high income may be inclined to lobby for public expenditures targeted to universities or cultural facilities just in order to prevent public money from being spent on transfers.

The main results of the analysis are as follows: First, the differential treatment of "rich" and "poor" individuals by the tax system implies that these individuals differ in their assessment of the desirability of public good provision – even if they derive the same utility from the public good. If tax revenues are used for a public good, this is more painful for low-skilled individuals because it necessitates a reduction of direct

[1] Examples are Boadway and Keen (1993), Nava et al. (1996), Gaube (2005) or Gahvari (2006). More generally, any study in public finance that derives a version of the Samuelson in the presence of distortionary taxation, imposes the assumption of a known distribution of preferences. See, for instance, Atkinson and Stern (1974), Wilson (1991), Sandmo (1998), Gaube (2000) or Hellwig (2004).

income transfers.
Second, whether or not this leads to an incentive problem depends on which dimension of heterogeneity among individuals is more pronounced. If public goods preferences are very intense, then two individuals who realize the same utility gain if a public good is provided but differ in their tax payments because of different earning abilities, have essentially the same views on the desirability of public good provision. In this case, the policy maker can easily acquire information on the distribution of preferences. To illustrate this, think of a measure of pollution control. As long as support or opposition to this policy measure is not primarily expressed by individuals with a high or a low earning ability, the policymaker is able to observe how the economy as a whole would be affected by this policy and hence be able to conduct an appropriate cost-benefit analysis.
If, by contrast, heterogeneity in earning abilities is very pronounced, then differences in the burden of taxation – as opposed to differences in public good preferences – determine whether an individual is made better off if a public good is provided. In this case, the "rich" want to have the public good even if their valuation of the public good is low, because it serves an instrument to limit the extent of redistribution. Likewise, less able individuals understate their preferences in order to avoid a cut of transfers. Hence, an informed decision on public good provision necessitates an adjustment of the transfer system that corrects these biases.
To eliminate these biases public good provision has to be made more costly for the "rich". Consequently, the optimal allocation is such that redistribution is increased whenever the public good is provided. Analogously, there is less redistribution, if the public good is not provided. Relative to an optimal income tax in the sense of Mirrlees (1971), which focusses solely on earning ability, income transfers are higher whenever a public good is provided and lower otherwise.

The model assumes that individuals either have a low or a high earning ability. Likewise, valuations of the public good are either high or low. Moreover, individuals have private information about these characteristics. For the economy as a whole, there is a problem of *information aggregation* because the distribution of public goods preferences among high and low-skilled individuals, respectively, is a priori unknown.[2]
In the attempt to link the theory of optimal income taxation with the literature on the free-rider problem in public good provision one must deal with an additional conceptual problem. The former analyzes economies that are "large", in the sense that no single individual has an impact on the tax system or the decision whether or not to provide a public good. The literature on the free-rider problem, by contrast, is concerned with an economy that is "small" in the sense that any one individual exerts a noticeable influence on public good provision. In this paper, the economy is assumed to be large

[2]This assumption of aggregate uncertainty distinguishes this paper from the multidimensional screening problems in Armstrong and Rochet (1999), Cremer et al. (2001, 2003) or Hellwig (2004).

as in the theory of optimal income taxation.[3] Moreover, *interest groups* – as opposed to single individuals – can affect spending decisions and a requirement of collective incentive compatibility is imposed.[4] In the model, an *interest group* is any group of individuals who have an incentive to coordinate their behavior in such a way that the policymaker ends up with a wrong perspective on the desirability of public good provision. The existence of such interest groups is not exogenously assumed. Interests groups arise endogenously if the tax system makes manipulations of public expenditure choices attractive.

In the setup of this paper, the relevant interest groups are easily characterized. In order to ensure that the preferences for a public good are communicated in a non-manipulative way, one has to make sure that neither an interest group which is supported only by high-skilled individuals nor an interest group which is supported only by low-skilled individuals has an incentive to misrepresent preferences for the public good.

The remainder of the paper is organized as follows. Section 2 defines the environment. Section 3 derives preferences for a public good as a function of an optimal income tax system. Section 4 contains the definition of an income tax that is non-manipulable by interest groups. In Section 5 the *optimal* non-manipulable income tax is characterized. The last section contains concluding remarks. All proofs are in the appendix.

2 The environment

The economy consists of a continuum of individuals. The preferences of individual j are represented by the utility function

$$\theta^j Q + u(C) - v\left(\frac{Y}{w^j}\right) . \tag{1}$$

$Q \in \{0,1\}$ stands for a public project and θ^j is a taste parameter. C denotes consumption of private goods, and Y denotes the agent's contribution to the economy's output. w^j is a productivity parameter that can be interpreted as a wage rate and $L = \frac{Y}{w^j}$ is the number of hours worked. The functions u and v are strictly increasing and twice continuously differentiable.

The characteristics of individual j, (w^j, θ^j), are private information of individual j. For the economy as a whole, these characteristics are the realization of a random variable $(\tilde{w}^j, \tilde{\theta}^j)$. The productivity parameter $\tilde{w}^j$ takes the values w_1 and w_2, $w_2 > w_1$, with probability $\frac{1}{2}$ each. By a law of large numbers,[5] these probabilities also indicate the population shares of people with these productivity levels, with no aggregate uncertainty. The cross section distribution of the taste parameter $\tilde{\theta}^j$ does involve aggregate

[3] This has the implication that the *Taxation Principle*, established by Hammond (1979) and Guesnerie (1995), is applicable. Hence, an optimal income tax is an optimal screening mechanism.

[4] This approach is discussed in more detail in Bierbrauer (2007). It uses ideas from the literature on mechanism design problems under a threat of collusion among agents; see Laffont and Martimort (1997, 1999).

[5] See Judd (1985) or Al-Najjar (2004).

uncertainty. $\tilde{\theta}^j$ takes the values θ_L and θ_H, $\theta_L > \theta_H$, and is correlated with $\tilde{w}^j$ and with a variable $\tilde{s}$ which I refer to as the *state* of the economy. The latter variable takes four possible values, s_{LL}, s_{LH}, s_{HL} and s_{HH}. It affects the probability of $\tilde{\theta}^j = \theta_H$ conditional on $\tilde{w}^j$, as summarized in the following table

	s_{LL}	s_{LH}	s_{HL}	s_{HH}
$\rho_1(s_{xy})$	α_1	α_1	β_1	β_1
$\rho_2(s_{xy})$	α_2	β_2	α_2	β_2

where $\rho_t(s_{xy}) := \text{prob}(\tilde{\theta}^j = \theta_H \mid \tilde{w}^j = w_t, \tilde{s} = s_{xy})$, for $t \in \{1, 2\}$. It is assumed that $\alpha_1 < \beta_1$ and $\alpha_2 < \beta_2$.

Due to the law of large numbers these conditional probabilities are interpreted as the fraction of low-skilled and high-skilled individuals, respectively, who have a high taste parameter. Whenever $x = L$, then the state s_{xy} is such that within the group of individuals who have a low skill level only a small fraction α_1 has a high taste parameter. By contrast, if $x = H$ then a large share β_1 of the low-skilled has a high taste parameter. Analogously, $y = L$ indicates that only a small subset of the high-skilled has a high taste parameter, whereas $y = H$ indicates that a large subset of this group has a high taste parameter.

The characterization of an optimal rule for both income taxation and public good provision is treated as a problem of mechanism design. A *social choice function* specifies for each state s a public good provision level $Q(s)$, a consumption level $C(w, \theta, s)$ and an output requirement $Y(w, \theta, s)$, for each (w, θ) in the set of characteristics $\Gamma := \{\theta_L, \theta_H\} \times \{w_1, w_2\}$. Attention is limited to social choice functions that can be reached via an anonymous allocation mechanism. Such an allocation mechansism consists of an abstract message space and outcome functions such that (i) the decision on public good provision is a function of the distribution of messages and (ii) individual i's consumption and output depend on i's message and the distribution of messages.[6] According to the Revelation Principle,[7] it entails no loss of generality to limit attention to direct mechanisms, i.e. to mechanism where an individual message consists of an announced taste parameter and an announced productivity level. A social choice function can be reached by such a mechanism if and only if it is *individually incentive compatible* and *feasible*. In the following, such a social choice functions is called *admissible*.

Feasibility requires that, in every state s, aggregate consumption, including the cost of public good provision, must not fall short of aggregate production,

$$\begin{aligned}&\frac{1}{2}\sum_{t=1}^{2}(1-\rho_t(s))Y(w_t,\theta_L,s)+\rho_t(s)Y(w_t,\theta_H,s)\\ &\geq kQ(s)+\frac{1}{2}\sum_{t=1}^{2}(1-\rho_t(s))C(w_t,\theta_L,s)+\rho_t(s)C(w_t,\theta_H,s)\ ,\end{aligned}$$

[6]Such an allocation mechanism is called anonymous because a permutation of messages does neither affect the decision on public good provision nor the consumption-income combination that is assigned to a specific message.

[7]A version of the Revelation Principle that applies to this setting is proven in Bierbrauer (2007).

where k denotes the aggregate resource requirement of public good provision.
Individual incentive compatibility requires that, in every state s, an individual with characteristics (w, θ) prefers the consumption-output combination dedicated to him over any other alternative combination, i.e. for every (w, θ) and every $(\hat{w}, \hat{\theta})$,

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

These incentive compatibility conditions require that a truthful announcement of characteristics is a best response from an *ex post* perspective; that is, after the resolution of uncertainty about the state of the economy.[8]
Bierbrauer (2007) contains a generalization of the *Taxation Principle*, due to Hammond (1979) and Guesnerie (1995). To illustrate this result, say that an allocation can be reached using an income tax if there exists a function $T : \mathbb{R}_+ \times S \to \mathbb{R}$ such that, in every state s, an individual's consumption-income pair solves

$$\max \quad u(C) - v\left(\frac{Y}{w}\right) \quad \text{s.t.} \quad C = Y - T(Y,s)$$

and, moreover, aggregate tax revenues are no less than the cost of public good provision. The *Taxation Principle* states that an allocation is admissible if and only if it is decentralizable by means of an income tax. Ex post incentive compatibility is hence equivalent to the requirement that individuals know the tax system that relates pre-tax and after-tax income when confronted with a consumption-leisure trade-off.
The provision rule Q does not enter the incentive compatibility conditions. 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.

The mechanism designer chooses a social choice function in order to maximize expected utilitarian welfare

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

where $W(s)$ denotes welfare in state s, and the mechanism designer's prior beliefs are given by $\pi_{LL} := \text{prob}(\tilde{s} = s_{LL})$, $\pi_{LH} := \text{prob}(\tilde{s} = s_{LH})$, etc. The mechanism design problem is simplified by the following lemma.

Lemma 1 A social choice function is admissible if and only if, in every state s, the following two properties hold:

i) *Equal treatment of tastes*:

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

for every productivity level w_t and every pair of taste parameters θ and θ'.

[8] Ex post incentive compatibility implies, in particular, that no assumptions on individual beliefs about the liklelihood of different states are needed. For a more detailed discussion see, e.g. Bergemann and Morris (2005), Kalai (2004) or Chung and Ely (2003).

ii) *Revelation of skill levels*:

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

for every taste parameter θ and every pair of productivity levels w_t and $w_{t'}$.

The provision rule for the public good $Q(\cdot)$ does not affect individual incentives. As a consequence, the mechanism designer has to specify the allocation of private goods in such a way that an individual with characteristics (w_t, θ_L) is indifferent between the bundle $(C(w_t,\theta_L,s), Y(w_t,\theta_L,s))$ and the bundle $(C(w_t,\theta_H,s), Y(w_t,\theta_H,s))$. Efficiency requires that this joint utiliy level is generated with minimal resources. Hence, it is without loss of generality to assume that

$$(C(w_t,\theta,s), Y(w_t,\theta,s)) = (C(w_t,\theta',s), Y(w_t,\theta',s)) ;$$

i.e. individuals who differ only in their taste parameter end up with the same (C,Y)-combination.

Given this observation, one can write $(C_t(s), Y_t(s))$ instead of $(C(w_t,\theta_L,s), Y(w_t,\theta_L,s))$ and $(C(w_t,\theta_H,s), Y(w_t,\theta_H,s))$. With this notation, utilitarian welfare in state s becomes

$$W(s) \quad := \quad \bar{\theta}(s)Q(s) + \tfrac{1}{2}\sum_{t=1}^{2}\left[u(C_t(s)) - v\left(\tfrac{Y_t(s)}{w_t}\right)\right] ,$$

where $\bar{\theta}(s) := \frac{1}{2}\sum_{t=1}^{2}(1-\rho_t(s))\theta_L + \rho_t(s)\theta_H$ is the average of all individual taste parameters in state s. The feasibility constraint can be written as

$$\tfrac{1}{2}Y_1(s) + \tfrac{1}{2}Y_2(s) \quad \geq \quad kQ(s) + \tfrac{1}{2}C_1(s) + \tfrac{1}{2}C_2(s) , \tag{2}$$

and the constraint that individuals reveal their skills becomes: for all t and $t' \neq t$,

$$u(C_t(s)) - v\left(\tfrac{Y_t(s)}{w_t}\right) \quad \geq \quad u(C_{t'}(s)) - v\left(\tfrac{Y_{t'}(s)}{w_t}\right) . \tag{3}$$

3 Optimal admissible social choice functions

The theory of optimal income taxation focuses on the incentive problems that arise if individuals have private information on their earning abilities. As a benchmark, this section characterizes the optimal social choice function for this environment.

The problem is to choose $\{Q(s), Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s\in S}$ in order to maximize EW subject to the feasibility constraints in (2) and the incentive constraints in (3). The total utility that is realized by an individual j with productivity w_t in state s, at an optimal allocation, can be written as $\theta^j Q(s) + U_t(kQ(s))$, because $C_t(s)$ and $Y_t(s)$ depend only on the revenue requirement $kQ(s)$ in state s.[9]

The optimal provision rule requires to provide the public good whenever the average

[9] To be more precise, whenever there are two states s and s' such that $Q(s) = Q(s')$, then, for all t, $C_t(s) = C_t(s')$ and $Y_t(s) = Y_t(s')$.

utility gain from public good provision exceeds the average utility loss that results from the increase in the revenue requirement,

$$Q(s) = 1 \iff \bar{\theta}(s) \;\geq\; \tfrac{1}{2}\sum_{t=1}^{2} U_t(0) - U_t(k)\;.$$

This provision rule depends on the parameter values θ_L, θ_H, α_1, β_1, α_2 and β_2. To avoid a lengthy discussion of each conceivable case, I focus on one particular constellation.[10]

Assumption 1 The optimal provision rule, $Q^*(\cdot)$, requires to provide the public good in all states except s_{LL}:

$$\min\{\bar{\theta}(s_{LH}), \bar{\theta}(s_{HL})\} \;\geq\; \tfrac{1}{2}\sum_{t=1}^{2} U_t(0) - U_t(k) \;\geq\; \bar{\theta}(s_{LL})\;.$$

Assumption 2 The function $v(\cdot)$ satisfies

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

Proposition 1 Suppose Assumption 2 holds. Then:

$$U_1(k) - U_1(0) \;>\; U_2(k) - U_2(0) \;>\; 0\;.$$

Assumption 2 is a sufficient condition which implies that for less productive individuals the utility loss is larger if public good provision necessitates an increase of the revenue requirement. This assumption is imposed henceforth without further mention.[11] In more technical terms, the Proposition establishes a property of *decreasing differences* according to which a lower productivity level translates into a larger utility loss.

The logic behind this observation is as follows. Due to the concavity of u, a utilitarian planner would love to give all individuals the same consumption. At the same time he wants the more productive individuals to work harder because they suffer less from the necessity to generate income. Since individuals are privately informed about their skills, this outcome is out of reach, and a binding incentive constraint prevents the planner from extracting larger tax payments from the more productive. This implies that, whenever resources for the public good are needed, higher taxes for the "rich" are not incentive feasible, unless the taxes of the "poor" have been increased already. Consequently, public good provision is more harmful for less productive individuals.

There are different parameter constellations of the model according to whether taste parameters or skills are the main determinant of an individual's preference for the

[10] The assumption serves only expositional purposes. Alternative parameter specifications would not affect the logic of the analysis.

[11] An alternative assumption, which would also yield the result of Proposition 1, is that the function v is linear, see Weymark (1986) or Boadway et al. (2000).

public good. Assumption 1 and Proposition 1 imply that the following scenarios are possible.

$$\text{Sc.1:} \quad \theta_H \;\geq\; U_1(0) - U_1(k) \;>\; U_2(0) - U_2(k) \;\geq\; \theta_L \;,$$

$$\text{Sc.2:} \quad \theta_H \;\geq\; U_1(0) - U_1(k) \;\geq\; \theta_L \;>\; U_2(0) - U_2(k) \;,$$

$$\text{Sc.3:} \quad U_1(0) - U_1(k) \;>\; \theta_H \;>\; \theta_L \;>\; U_2(0) - U_2(k) \;.$$

Scenario 1: An individual's attitude to public good provision depends only on the taste parameter. Both less productive and more productive individuals benefit form public good provision if and only if their taste parameter is high. The public good thus has supporters of all skill levels and opponents of all skill levels. As an example, think of a hospital that is valued by all individuals who are in bad health – even if they are treated differently by the tax system due to differences in income.
Scenario 3: This is the opposite of Scenario 1 in the sense that now the heterogeneity in skill levels dominates the heterogeneity in taste parameters. The fact that the public good is crowding out redistribution now implies that high-skilled individuals want public good provision even when their taste parameter is low, while low-skilled individuals oppose the public good even when their taste parameter is high. To illustrate this think of a public opera house. If Scenario 3 applies, this public project is valued by all individuals with a high income. Even if they do not go to the opera very often, they still prefer this opportunity over an increase in direct income transfers. By contrast, from the perspective of less productive individuals, this public good is too expensive once the accompanying reduction in income transfers is taken into account.
Scenario 2: Scenario 2 is an intermediate case. For less productive individuals, the taste parameter determines whether or not they would benefit from public good provision. Productive individuals who don't suffer as much if the revenue requirement goes up desire public good provision even if their taste parameter is low. Consequently, this public good is valued by all individuals with a high income but only by some individuals with a low income. An example for this scenario can be found in the domain of environmental policy. Suppose a concern for clean air is shared by all individuals with an above average income and a corresponding measure of pollution control is supported by these agents. Among poor households, by contrast, this policy preference is not shared unanimously because it implies a more drastic reduction in net incomes.

4 Collective Incentive Compatibility and Interest groups

The preceding analysis was based on the assumption that the distribution of preferences for the public good among high-skilled and low-skilled individuals, respectively, is observable by a policymaker. While she is constraint by the fact that individuals have private information on their earning abilities, the state s of the economy is a treated as known.

Within the model, this assumption is problematic. To see this, suppose that *Scenario 2* applies, i.e. all high-skilled individuals want to have the public good. Moreover, if these individuals convince the mechanism designer that, among them, the share of individuals with a high taste parameter is high, $\rho_2(s) = \beta_2$, they ensure that the public good is provided. Moreover, there is an obvious way to achieve this: a collective lie of $\beta_2 - \alpha_2$ individuals whose true taste parameter equals θ_L, whenever $\rho_2(s) = \alpha_2$. Such a collective lie is not prevented by individual incentive compatibility. Due to the equal treatment of individuals who differ only in their taste parameters, each individual is willing to announce any taste parameter. Hence, a collective lie involving taste parameters is not undermined by individual incentive compatibility.
Moreover, the assumption of a known distribution of preferences is unsatisfactory if confronted with the ways in which the demand for public goods arises in reality. It is very common that interest groups try to convince politicians that certain public goods need to be provided.[12] These interest groups do not articulate the true preferences of their constituency. Rather, they seek to induce a certain political outcome. In the following I will thus formalize the view that the acquisition of information on preferences requires appropriate incentives.

A social choice function is said to be *collectively incentive compatible* if, for every state s, there does not exist a group of individuals J with true characteristics $\{(w^j, \theta^j)\}_{j \in J}$ and an alternative profile of characteristics $\{(\hat{w}^j, \hat{\theta}^j)\}_{j \in J}$ such that the following three properties hold simultaneously:

i) If the true profile $\{(w^j, \theta^j)\}_{j \in J}$ is replaced by $\{(\hat{w}^j, \hat{\theta}^j)\}_{j \in J}$ then the cross-section distribution of characteristics corresponds to state $\hat{s}$.

ii) *Unanimity.* In state $\hat{s}$ all members of J are strictly better off; for every $j \in J$,

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

iii) *Stability.* In state $\hat{s}$, any individual $j \in J$ with characteristics (w^j, θ^j) is indifferent between $(C(w^j, \theta^j, \hat{s}), Y(w^j, \theta^j, \hat{s}))$ and $(C(\hat{w}^j, \hat{\theta}^j, \hat{s}), Y(\hat{w}^j, \hat{\theta}^j, \hat{s}))$.

These conditions can equivalently be interpreted as equilibrium conditions for a direct revelation mechanism.[13] Collective incentive compatibility holds if, in every state s, truth-telling is a best response from the perspective of each subset of individuals under the assumption that all other individuals tell the truth; i.e. truth-telling is collectively

[12] The ubiquitousness of interests groups in the political process is demonstrated in Chapter 1 of Grossman and Helpman (2001).

[13] Bierbrauer (2007) proves a revelation principle for *coalition-proof implementation* in an ex post equilibrium. Accordingly, collective incentive compatibility is a necessary condition for the implementability of a social choice function if individuals can collude.

a best response from an ex post perspective. Once the state of the economy has been revealed, no subset of agents would want to revise their announcements.[14]

The stability requirement implies that the announcements of deviating individuals are individually best responses. Recall that, by individual incentive compatibility, an individual with characteristics (w^j, θ^j) weakly prefers $(C(w^j, \theta^j, \hat{s}), Y(w^j, \theta^j, \hat{s}))$ over $(C(\hat{w}^j, \hat{\theta}^j, \hat{s}), Y(\hat{w}^j, \hat{\theta}^j, \hat{s}))$. If this preference was strict, then an individual who benefits from the collective deviation would be tempted to free-ride on coalition formation; that is, to take the outcome $\hat{s}$ as given and to reveal her characteristics truthfully. Imposing a requirement of stability follows Laffont and Martimort (1997, 2000) who argue that the formation of an interest group is confronted with the same incentive problems as the general mechanism.

The specific assumptions about the information structure of the economy imply that the requirement of collective incentive compatibility can be simplified considerably. In the following I will show that the relevant interest groups are the ones that form along the dimension of skill heterogeneity. Interest groups either consist entirely of "rich" individuals or they consist entirely of "poor" individuals.

Proposition 2 A social choice function is collectively incentive compatible only if it satisfies the following conditions for a *collective revelation of tastes*:

i) There is no scope for collective manipulations of taste parameters by low skilled individuals. Equivalently, for all $x, \hat{x} \in \{L, H\}$ and all $y \in \{L, H\}$,

$$\theta_x Q(s_{xy}) + V_1(s_{xy}) \quad \geq \quad \theta_x Q(s_{\hat{x}y}) + V_1(s_{\hat{x}y}) \; , \tag{4}$$

ii) There is no scope for collective manipulations of taste parameters by high skilled individuals. Equivalently, for all $y, \hat{y} \in \{L, H\}$ and all $x \in \{L, H\}$,

$$\theta_y Q(s_{xy}) + V_2(s_{xy}) \quad \geq \quad \theta_y Q(s_{x\hat{y}}) + V_2(s_{x\hat{y}}) \; , \tag{5}$$

where $V_t(s) := u(C_t(s)) - v(Y_t(s)/w_t)$ is a shorthand for the utility that individuals with skill level w_t derive in state s from their consumption-income combination.

The assumption that the cross-section distribution of skills is the same in every state s implies that collective deviations that involve false announcements of skill levels are superfluous.[15] Hence, it entails no loss of generality to assume that interest groups

[14] An alternative approach is considered by Bierbrauer and Hellwig (2007). In that paper, deviating individuals first report to a fictitious coalition organizer who coordinates the actions of coalition members such that they are made better off in expectation. If one combines this notion of coalition-proofness with a requirement of robustness with respect to assumptions on prior beliefs, see Bergemann and Morris (2005), then this implies the collective incentive constraints that arise in this paper.

[15] Whenever a group J can change the perceived state from state s to state $\hat{s}$ via some profile $\{(\hat{w}^j, \hat{\theta}^j)\}_{j \in J}$, then it can also do so by means of a profile that involves truthful announcements of skill levels.

manipulate only the distribution of announced taste parameters.
The collective incentive condition in (4) addresses manipulations by low-skilled individuals. Suppose the true state of the economy is s_{Ly} and consider a collective deviation that prescribes a false taste announcement for a measure of $\beta_1 - \alpha_1$ low-skilled individuals with a low taste parameter – and that thereby induces an announced state $\hat{s} = s_{Hy}$. This manipulation is stable because it prescribes only false announcements of taste parameters and individuals who differ only in their taste parameters are treated equally. Hence, from an individual's perspective, announcing a false taste parameter yields the same payoff as a truthful announcement. Consequently, this manipulation is eliminated only if it is not attractive for low skilled individuals, $\theta_L Q(s_{Ly}) + V_1(s_{Ly}) \geq \theta_L Q(s_{Hy}) + V_1(s_{Hy})$. Analogously, the constraint in (5) rules out manipulations by individuals who are high-skilled.

In addition to (4) and (5), collective incentive compatibility requires that there is no scope for interest groups that mislead the mechanism designer, both with respect to distribution of preferences among the high-skilled and with respect to the distribution of preferences among the low-skilled. However, the following proposition establishes that this constraint will never be binding.

Proposition 3 Consider a social choice function that maximizes *EW* subject to admissibility, (4) and (5). Then, for any state s, there is no collective manipulation that makes low-skilled and high-skilled individuals better off.

If the mechanism designer chose a social choice function such that both high-skilled and low-skilled individuals would like to correct the decision on public good provision, then it must be the case that the initial allocation has not been welfare maximizing. Consequently, only interest groups that are homogeneous with respect to the skills of their members are a relevant constraint for the mechanism designer.

The following Proposition shows that collective incentive compatibility is a genuine constraint: The optimal social choice function from Section 2 satisfies the collective incentive conditions only in Scenario 1.

Proposition 4 The social choice function that maximizes *EW* subject to admissibility satisfies (4) and (5) if and only if *Scenario 1* is given.

Scenario 1 is defined by the property that, at the optimal admissible allocation, all individuals with a low taste parameter oppose public good provision, while all individuals with a high taste parameter desire provision, irrespective of their skill level. The Proposition states that this is the only case in which the mechanism designer gets the information on public goods preferences for free. In all other case, collective incentive

constraints necessitate a deviation from the optimal admissible allocation.
The provision rule Q^*, that is part of the optimal admissible social choice function, is increasing in the population share of individuals with a high taste parameter. Hence, a collective lie of individuals with a high taste parameter would reduce the set of states in which the public good is provided. But since in *Scenario 1* all individuals with a high taste parameter are better off in states with $Q = 1$, this collective deviation is not attractive. Likewise, all individuals with a low taste parameter want to avoid provision and hence do not deviate from the truth. Under *Scenarios 2* and *3*, this property is violated.

5 Implications of collective incentive compatibility

It has been shown in Proposition 4 that the optimal admissible allocation triggers the formation of manipulative interest groups in Scenarios 2 and 3. This section characterizes the optimal allocation that satisfies not only feasibility and individual incentive compatibility but also the collective incentive constraints in (4) and (5).
I will focus on Scenario 2.[16] This implies that the optimal admissible allocation violates constraint (5). A collective incentive problem arises because high-skilled individuals suffer so little from an increase in the revenue requirement that they want public good provision even if they have a low taste parameter. In state s_{LL}, these individuals exaggerate their preferences in order to convince the mechanism designer that the true state is s_{LH} and that the public good should be provided.
The mechanism designer can respond in two ways. He can either stick to provision rule Q^*, which is part of the optimal admissible social choice function, and distort the accompanying tax system in such a way that the incentive to exaggerate is eliminated. Alternatively, he can distort the provision rule for the public good. For instance, if the mechanism designer chooses to implement a provision rule which prescribes to install the public good if and only if a large fraction of the low-skilled has a high valuation of the public good, $Q(s) = 1 \iff \rho_1(s) = \beta_1$, then there is no need to acquire information on the distribution of preferences among the high-skilled. This would also eliminate the collective incentive problem.
How the mechanism designer combines these two remedies depends on the intensity of the collective incentive problem. Suppose first that the mechanism designer sticks to provision rule Q^*. Then, to ensure that high-skilled individuals do not exaggerate, he has to make states with $Q = 1$ less attractive. Consequently, there is more redistribution whenever $Q = 1$ and less redistribution whenever $Q = 0$. If a "small" adjustment of the transfer scheme suffices to fix the collective incentive problem, then a choice of provision rule Q^* may still be still optimal. If by contrast a "large" distortion of the transfer system is needed, there will be a new collective incentive problem. Low-skilled individuals will start to exaggerate their preferences for the public good because the increase in redistribution has made states with $Q = 1$ more attractive for them. An

[16] At the end of this section I discuss how the results have to modified if Scenario 3 applies.

implementation of Q^* then yields two binding collective incentive constraints, hence a very substantive distortion of the income tay system.
To formalize these arguments I will first solve for the optimal allocation taking the provision rule Q^* as given. This exercise provides an answer to the question how the system of income transfers has to be distorted, whenever public good provision relies on information about the preferences of the "rich". I will then analyze under what circumstances the choice of such a provision rule is optimal.

5.1 The optimal income tax, taking Q^* as given

The following Lemma characterizes the Pareto-frontier in a neighborhood of the optimal admissible allocation. It is the main tool for a characterization of the optimum if, in addition, the collective incentive constraints (4) and (5) are taken in into account.

Lemma 2 Denote by $V_1^*(\bar{V}_2, r)$ the value function of the following problem:

$$\begin{array}{lll} \max_{C_1,Y_1,C_2,Y_2} & u(C_1) - v\left(\frac{Y_1}{w_1}\right) & \\ \text{s.t.} & \frac{1}{2}[Y_1 - C_1] + \frac{1}{2}[Y_2 - C_2] \geq r & (\text{BC})\ , \\ & u(C_1) - v\left(\frac{Y_1}{w_2}\right) \leq \bar{V}_2 & (\text{RS}_2)\ , \\ & u(C_2) - v\left(\frac{Y_2}{w_2}\right) = \bar{V}_2\ . & \end{array} \tag{6}$$

For all r, V_1^* is a continuous and strictly concave function of $\bar{V}_2$ with a unique maximum. For $\bar{V}_2 = U_2(r)$ – i.e. at the *optimal admissible allocation* with a revenue requirement of r – V_1^* is strictly decreasing in $\bar{V}_2$.

Lemma 2 shows that there is a well defined range of parameters for which there is indeed a tradeoff between the utility of the "rich" and the utility of the "poor".[17] Moreover, the utilitarian optimum does not lie at the boundary of the region where the tradeoff prevails; i.e. a utilitarian mechanism designer does not maximize redistribution.

[17] This is not trivial as there is a region where low-skilled individuals 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 for by the utility loss from a more severe *distortion at the bottom*. See the appendix for a mathematical formulation.

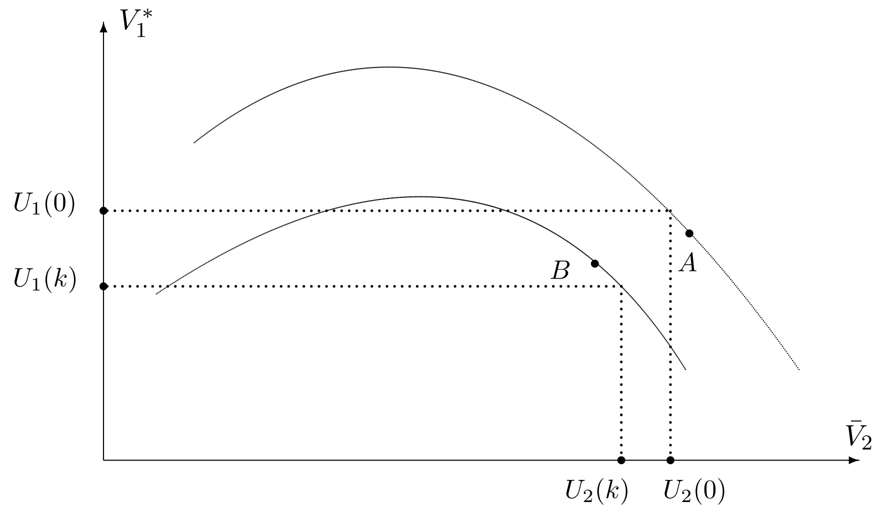

Figure 1: The graph shows the Pareto frontier of the set of admissible allocations for $r = 0$ and $r = k$, respectively. At the optimal admissible allocation, collective incentive compatibility fails because $U_2(0) - U_2(k)$ is smaller than θ_L.

Proposition 5 Suppose provision rule Q^* is taken as given. Consider the problem of maximizing EW subject to admissibility and the collective incentive constraints (4) and (5). Denote by $\{V_t^{**}(s)\}_{s \in S}$ the utility levels realized by individuals with skill level w_t at a solution to this problem. There exists $\bar{\theta}_L$ such that if $\theta_L \leq \bar{\theta}_L$, then:

i) Whenever the public good is not installed, then, relative to the optimal admissible allocation, low-skilled individuals are worse off and high-skilled individuals are better off. Formally, $V_1^{**}(s_{LL}) < U_1(0)$ and $V_2^{**}(s_{LL}) > U_2(0)$.

ii) In all states with public good provision, individuals derive the same utility from their consumption-income combination. Formally, for all $t \in \{1, 2\}$, $V_t^{**}(s_{LH}) = V_t^{**}(s_{HL}) = V_t^{**}(s_{HH})$.

iii) Whenever the public good is installed, then, relative to the optimal admissible allocation, low-skilled individuals are better off and high-skilled individuals are worse off. Formally, let s be such that $Q(s) = 1$, then $V_1^{**}(s) > U_1(k)$ and $V_2^{**}(s) < U_2(k)$.

iv) The following collective incentive constraint for high-skilled individuals is binding, $V_2^{**}(s_{LL}) - V_2^{**}(s_{LH}) = \theta_L$. The collective incentive constraints for low-skilled individuals, $\theta_H \geq V_1^{**}(s_{LL}) - V_1^{**}(s_{HL}) \geq \theta_L$, are not binding.

Given Scenario 2, the constraint $V_2^{**}(s_{LL}) - V_2^{**}(s_{LH}) \geq \theta_L$ is violated at the optimal admissible allocation, implying that high-skilled individuals exaggerate their preferences for the public good. Proposition 5 characterizes the optimal allocation with the property that this is constraint is binding.

The mechanism designer deviates along the Pareto-frontier such that, from the perspective of the "rich", public good provision becomes less attractive. An optimal choice of this deviation displays an increased level of redistribution in states with public good provision and less redistribution in states with non-provision: In state s_{LL} the public good is not provided and high-skilled individuals receive a C-Y pair that generates a utility level above $U_2(0)$. In all other states, the public good is provided and those individuals get a C-Y pair that implies a utility level below $U_2(k)$.
Given this adjustment, low-skilled individuals are made as well off as possible. In particular, this implies that less productive individuals can be made better off in states with public good provision. As high-skilled individuals receive a utility level below $U_2(k)$, this leaves room to raise the utility of low-skilled individuals above $U_1(k)$. Analogously, in states without public good provision, low-skilled individuals are worse off. As the utility level of the "rich" exceeds $U_2(0)$, a utility level of $U_1(0)$ is out of reach for the "poor".
Proposition 5 is based on the assumption that θ_L is not "too high". The assumption $\theta_L \leq \bar{\theta}_L$ ensures that the collective incentive constraint,

$$V_1^{**}(s_{LL}) - V_1^{**}(s_{HL}) \quad \geq \quad \theta_L \ , \tag{7}$$

for low-skilled individuals is not binding. The optimal allocation in Proposition 5 is such that the utility difference $V_1(s_{LL}) - V_1(s_{HL})$ is smaller than $U_1(0) - U_1(k)$. If the parameter θ_L is small, then $V_1(s_{LL}) - V_1(s_{HL})$ can indeed be decreased without violating the incentive constraint in (7). However, if θ_L is too large, the combined effect of public good provision and more redistribution in state s_{LH} implies that low-skilled individuals have an incentive to exaggerate their preferences in state s_{LL}. An implementation of provision rule Q^* then implies that there are two binding collective incentive constraints,

$$V_1^{**}(s_{LL}) - V_1^{**}(s_{HL}) = \theta_L \quad \text{and} \quad V_2^{**}(s_{LL}) - V_1^{**}(s_{LH}) = \theta_L \ .$$

Below I will show that, whenever an implementation of Q^* yields two binding collective incentive constraints, then it is optimal to choose a different provision rule.
Based on these considerations, I will use the following terminology. Provision rule Q^* faces a *modest* collective incentive problem if the optimal tax system that implements Q^* implies that the constraint $V_2^{**}(s_{LL}) - V_2^{**}(s_{LH}) \geq \theta_L$ is binding and the constraint $V_1^{**}(s_{LL}) - V_1^{**}(s_{HL}) \geq \theta_L$ is slack. Q^* faces a *severe* collective incentive problem if these constraints are both binding.

5.2 The optimal provision rule

I now investigate whether provision rule Q^* is part of an optimal allocation. The necessity to distort the accompanying tax system may imply that a different provision rule turns out to be superior. One alternative is to provide the public good in every state of the world, $Q(s) = 1$, for all 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 acquire information. Hence, there is no need to distort the transfer system.
If Q^* faces a *modest* incentive problem, it depends on the planner's prior whether or not Q^* is the optimal provision rule. To see this, suppose first that p_{LL} is very small. Then, provision rule $Q(s) = 1$, for all s, leads to a larger welfare level than Q^* because the state in which a deviation from the optimal admissible allocation occurs is very unlikely. Alternatively, suppose that all states of the world are equally likely and suppose the parameter θ_L is such that Q^* requires only a "small" deviation from the optimal admissible allocation, i.e. θ_L is only slightly larger than $U_2(0) - U_2(k)$ – in terms of Figure 1, the points A and B are very close to the optimal admissible allocation. In this case, the adjustments of the transfer system required under Q^* are negligible in welfare terms, and Q^* is superior to any alternative provision rule.
In contrast, with a *severe* incentive problem, Q^* will not be chosen. To see this, suppose that the *C-RT* 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; that is, all individuals with a low taste parameter are indifferent between public good provision and non-provision. Hence, these individuals are indifferent between the provision rules Q^* and $Q(s) = 1$, for all s. However, all individuals with a high taste parameter prefer the latter provision rule and, moreover, it avoids any departure from $U_1(k)$ and $U_2(k)$. Hence, utilitarian welfare is higher in every state of the economy. These considerations are summarized in the following proposition.

Proposition 6 If Q^* faces a *modest* collective incentive problem, then it depends on the parameters of the model, whether or not Q^* is part of an optimal allocation. If Q^* faces a *severe* collective incentive problem, then Q^* is not part of an optimal allocation.

I do not characterize the optimal provision rule in more detail. In total their are six provision rules that are consistent with collective incentive compatibility.[18] A complete characterization of the optimum requires an analysis similar to the one for Q^* for each of these rules; that is, one would has to determine for each of them the pattern of binding collective incentive constraints and their welfare implications. For Scenario 2, the main results can be summarized as follows: if provision rule Q^* – or any other rule that makes the decision on public good provision dependent on the distribution of preferences among high-skilled individuals – is chosen, the planner has to accept the necessity of excessive redistribution if the public good is provided, and of suboptimal redistribution if not. However, this may imply that a different provision rule, which does not depend on $\rho_2(s)$, becomes preferable.

[18] Conditions (4) and (5) imply that the following monotonicity conditions have to be satisfied $Q(s_{LL}) \leq Q(s_{LH}) \leq Q(s_{HH})$ and $Q(s_{LL}) \leq Q(s_{HL}) \leq Q(s_{HH})$. This singles out the the following provision rules: (i) the constant provision rules $Q(s) = 1$ for all s and $Q(s) = 0$ for all s; (ii) the provision rules Q^1 defined by $Q(s_{xy}) = 1$ if and only if $x = H$ and the provision rule Q^2 defined by $Q(s_{xy}) = 1$ if and only if $y = H$; (iii) Q^* and the provision rule Q' defined by $Q(s) = 1 \Longleftrightarrow s = s_{LL}$.

5.3 Scenario 3

Under *Scenario 3* an individual's view on the desirability of public good provision is entirely determined by her earning ability. At the optimal admissible allocation, all high-skilled individuals are better off if the public good is provided and all low-skilled individuals are worse off, irrespective of their taste parameters. This implies that there are now two collective incentive problems. The first one is familiar from Scenario 2, high-skilled individuals exaggerate their preferences for the public good. In addition, low-skilled individuals want the mechanism designer to believe that $\rho_1(s) = \alpha_1$ even if $\rho_1(s) = \beta_1$ because their utility burden from paying for the public good exceeds θ_H.
Given provision rule Q^*, collective truth-telling of less productive individuals, requires that states with $Q = 1$ are made more attractive relative to states with $Q = 0$ and that the utility difference $V_1(s_{LL}) - V_1(s_{HL})$ is decreased relative to $U_1(0) - U_1(k)$. Simultaneously, for high-skilled individuals, the attractiveness of public good provision has to be decreased. Fortunately, these incentive corrections tend to complement each other: If the utility difference between provision and non-provision is decreased from the perspective of the low-skilled this implies that, from the perspective of the high-skilled, this difference is increased. Hence, under *Scenario 3*, provision rule Q^* may face an incentive problem that is *modest* in the sense that if the collective incentive constraint for high-skilled individuals is binding, this implies that the constraint for low-skilled individuals is slack. Otherwise, provision rule Q^* faces a *severe* incentive problem. Again, this distinction is decisive for the question wether Q^* can be part of an optimal allocation.

6 Concluding Remarks

The analysis has shown that, under an optimal income tax, an individual's earning ability affects his willingness to pay for public goods. Consequently, individuals may lobby for their most preferred expenditure policies in a way that makes it impossible to measure the costs and benefits of these policies. The attempt to take the manipulative impact of interest groups into account gives rise to a new set of incentive considerations. In general, they necessitate a distortion of the tax system or the provision rule for public goods. Whenever information acquisition is desirable in the presence of collective incentive constraints, then there is a complementarity between redistribution and public good provision: Public good provision implies an increase in redistribution.
This raises the question how to assess these deviations from a welfare perspective. As the analysis has shown, it is possible that such a deviation will make one group of individuals better off while hurting another group, i.e. there is no departure from constrained efficiency. However, these deviations place an additional welfare cost on redistribution. Consequently, one has to tradeoff the utilitarian welfare gains from a more favorable solution to the *equity-efficiency tradeoff* with those from a more favorable solution to the *free-rider problem.* Generally, a deviation from an *optimal*

income tax, as typically defined in the literature, is desirable in order to improve the possibility to acquire information on the distribution of preferences.

References

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

Armstrong, A. and Rochet, J. (1999). Multi-dimensional screening: A user's guide. *European Economic Review*, 43:959–979.

Atkinson, A. and Stern, N. (1974). Pigou, taxation and public goods. *Review of Economic Studies*, 41:119–128.

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

Bierbrauer, F. (2007). Robust and coalition-proof mechanisms for income taxation and public good provision. Work in Progress.

Bierbrauer, F. and Hellwig, M. (2007). Public good provision in a large economy. Work in progress.

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. (2003). Ex-post incentive compatible mechanism design. Mimeo, Northwestern University.

Cremer, H., Pestieau, P., and Rochet, J. (2001). Direct versus indirect taxation: The design of the tax structure revisited. *International Economic Review*, 42:781–799.

Cremer, H., Pestieau, P., and Rochet, J. (2003). Capital income taxation when inherited wealth is not observable. *Journal of Public Economics*, 87:2475–2490.

Gahvari, F. (2006). On the marginal costs of public funds and the optimal provision of public goods. *Journal of Public Economics*, 90:1251–1262.

Gaube, T. (2000). When do distortionary taxes reduce the optimal supply of public goods. *Journal of Public Economics*, 76:151–180.

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

Grossman, G. and Helpman, E. (2001). *Special Interest Politics*. Cambridge, MA, MIT Press.

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.

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.

Laffont, J. and Martimort, D. (2000). Mechanism design with collusion and correlation. *Econometrica*, 68:309–342.

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

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.

Sandmo, A. (1998). Redistribution and the marginal cost of public funds. *Journal of Public Economics*, 70:365–382.

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

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

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

A Appendix

Proof of Lemma 1

Claim 1. To prove the only if-part, note that, because preferences satisfy the separability property stated in equation (1), properties i) and ii) are implied by individual incentive compatibility. To prove the if-part, suppose a social choice function satisfies i) and ii) but is not incentive compatible. Then there exist (w, θ), $(\hat{w}, \hat{\theta})$ 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 i) and ii) one has:

$$\begin{aligned} & u(C(\hat{w},\hat{\theta},s)) - v\left(\frac{Y(\hat{w},\hat{\theta},s)}{w}\right) = u(C(\hat{w},\theta,s)) - v\left(\frac{Y(\hat{w},\theta,s)}{w}\right) \\ & \leq u(C(w,\theta,s)) - v\left(\frac{Y(w,\theta,s)}{w}\right) . \end{aligned}$$

Hence, a contradiction.

■

Proof of Proposition 1

It is helpful to introduce the following auxiliary problem, which does not include a public good but instead an exogenous revenue requirement $r \geq 0$.

$$\begin{array}{ll} \max_{C_1,Y_1,C_2,Y_2} & u(C_1) - v\left(\frac{Y_1}{w_1}\right) + u(C_2) - v\left(\frac{Y_2}{w_2}\right) \\ \text{s.t.} & Y_1 - C_1 + Y_2 - C_2 \geq r , \\ & u(C_1) - v\left(\frac{Y_1}{w_1}\right) \geq u(C_2) - v\left(\frac{Y_2}{w_1}\right) , \\ & u(C_2) - v\left(\frac{Y_2}{w_2}\right) \geq u(C_1) - v\left(\frac{Y_1}{w_2}\right) . \end{array} \tag{8}$$

Denote by $U_t(r)$ the utility level realized by individuals with productivity level w_t at a solution to problem (8).
As is well known in the literature (see e.g. Stiglitz (1982)), at a solution to problem (8) the feasibility constraint and the constraint $u(C_2) - v\left(\frac{Y_2}{w_2}\right) \geq u(C_1) - v\left(\frac{Y_1}{w_2}\right)$ are binding, implying that there is a *distortion at the bottom* and *no distortion at the top*:

$$MRS_1^* := \frac{\frac{1}{w_1} v'\left(\frac{Y_1^*(r)}{w_1}\right)}{u'(C_1^*(r))} < 1 \quad \text{and} \quad MRS_2^* := \frac{\frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right)}{u'(C_2^*(r))} = 1 , \tag{9}$$

where $(Y_1^*(r), C_1^*(r), Y_2^*(r), C_2^*(r))$ denotes the solution to problem (8), which depends on the revenue requirement r. A Lagrangean approach yields the following first order conditions:

$$\frac{u'(C_1^*(r))}{u'(C_2^*(r))} = \frac{(1 - MRS_1^*) + (1 - \widehat{MRS}^*)}{MRS_1^* - \widehat{MRS}^*} , \tag{10}$$

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

$$Y_1^*(r) - C_1^*(r) + Y_2^*(r) - C_2^*(r) = r , \tag{11}$$

$$u'(C_2^*(r)) = \frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right) , \tag{12}$$

$$u(C_1^*(r)) - v\left(\frac{Y_1^*(r)}{w_2}\right) = u(C_2^*(r)) - v\left(\frac{Y_2^*(r)}{w_2}\right) . \tag{13}$$

This system of equations 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.

Claim 1.

$$\frac{dY_1^*(r)}{dr} > 0,\ \frac{dC_1^*(r)}{dr} < 0 \quad \text{and} \quad \frac{dY_2^*(r)}{dr} > 0,\ \frac{dC_2^*(r)}{dr} < 0\ .$$

Proof. After lengthy calculations, one finds that

$$\frac{dY_1^*(r)}{dr} = \frac{\alpha\pi + \epsilon\gamma}{\beta\pi + \delta\epsilon + \rho} \tag{14}$$

where,

$$\alpha := \frac{2u'(C_2^*)}{u'(C_1^*) + u'(C_2^*)} > 0\ ,$$

$$\beta := \frac{\frac{1}{w_2}v'\left(\frac{Y_2^*}{w_2}\right) + u'(C_2^*)}{u'(C_1^*) + u'(C_2^*)} > 0\ ,$$

$$\gamma := \frac{2u'(C_1^*)}{u'(C_1^*) + u'(C_2^*)} \frac{\frac{1}{w_2^2}v''\left(\frac{Y_1^*}{w_2}\right)}{\frac{1}{w_2^2}v''\left(\frac{Y_1^*}{w_2}\right) - u''(C_2^*)} > 0\ ,$$

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

To see that $\delta > 0$, note that due the *distortion at the bottom* and $w_2 > w_1$,

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

This also implies that

$$\epsilon := -\frac{u''(C_2)}{2}\left[2u'(C_1) - \frac{1}{w_1}v'\left(\frac{Y_1^*}{w_1}\right) - \frac{1}{w_2}v'\left(\frac{Y_1^*}{w_2}\right)\right] > 0\ .$$

In addition,

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

To see that $\pi > 0$, note that equation (10) implies:

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

Finally,

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

Note that the premise of Proposition 1 implies that $\rho > 0$. Since all terms that appear on the right-hand-side of equation (14) are strictly positive it follows that Y_1^* increases in r. Moreover differentiating the first order conditions in (10) - (13) with respect to r also yields the following observations.

$$\frac{dC_1^*(r)}{dr} = -\alpha + \beta\frac{dY_1^*(r)}{dr} \quad \text{and} \quad \frac{dC_2^*(r)}{dr} = -\gamma + \delta\frac{dY_1^*(r)}{dr}\ .$$

The facts that $\gamma > \delta$ and that $\alpha > \beta$ imply that $\frac{dY_1^*(r)}{dr} < 1$ and hence that $\frac{dC_1^*(r)}{dr} < 0$ and that $\frac{dC_2^*(r)}{dr} < 0$. Finally, the fact that C_2^* decreases in r in conjunction with equation (12) implies that Y_2^* increases in r.

Claim 2.

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

Proof. From equation (13) it follows that:

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

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

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

To see also that the first inequality holds, note that, using (12), one has:

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

This expression is strictly negative by *Claim 1.*

■

Proof of Proposition 2

I show that the inequalities in (4) are both necessary and sufficient for the elimination of collective lies on taste parameters by any group of low skilled individuals.

I first note that the inequalities in (4) require that no mass $\beta_1 - \alpha_1$ of individuals with skill level w_1 is made better off by announcing a false taste parameter, taking as given that all individuals with skill level w_2, reveal their taste parameters.

The necessity of this condition follows from the fact that these manipulations induce by construction some state $\hat{s} \neq s$. They are also stable because individuals who differ only in their taste parameters are treated equally by individual incentive compatibility.

To prove the sufficiency of (4), I show that whenever there is some manipulating coalition, then there exists an *equivalent* manipulation which violates (4).

Let s be the actual state of the economy and consider an interest group J that induces $\hat{s}$. An interest group J' is said to be *equivalent* to J if it induces the same $\hat{s}$ and moreover has the same payoff consequence for every individual in the economy.

Note first that it is without loss of generality to assume that all individuals reveal their skills. Since the skill distribution is given, they could not induce a state $\hat{s}$ otherwise. Due to the information structure of the economy, $\hat{s}$ can be induced only if a mass $\beta_1 - \alpha_1$ of individuals reports a false taste parameter. Obviously, if a larger mass of individuals deviates, there is an equivalent manipulation where exactly a mass $\beta_1 - \alpha_1$ declares a false taste parameter.

■

Proof of Proposition 3

If attention is limited to allocations that satisfy (4) and (5), then an interest group with both high-skilled and low-skilled individuals can fulfill the unanimity requirement only if there exist two states s_{xy} and $s_{\hat{x}\hat{y}}$ with $\hat{x} \neq x$ and $\hat{y} \neq y$ such that both high-skilled and low-skilled individuals are better off if in state s_{xy} the outcome specified for state $s_{\hat{x}\hat{y}}$ is implemented.[19]

Conditions (4) and (5) imply that the following monotonicity condition of have to be satisfied

$$\begin{aligned} &Q(s_{LL}) \leq Q(s_{LH}), \quad Q(s_{LL}) \leq Q(s_{HL}) \ , \\ &Q(s_{LH}) \leq Q(s_{HH}), \quad Q(s_{HL}) \leq Q(s_{HH}) \ . \end{aligned}$$

This observation limits the number of admissible provision rules. *First*, there are the constant provision rules $Q(s) = 1$ for all s and $Q(s) = 0$ for all s. Obviously, the optimal accompanying choices of $C_t(s)$ and $Y_t(s)$ are also constant across states. This trivially implies that there is no incentive to manipulate the mechanism designer's perception of s.

Second, there are the provision rules Q^1 defined by $Q(s_{xy}) = 1$ if and only if $x = H$ and the provision rule Q^2 defined by $Q(s_{xy}) = 1$ if and only if $y = H$. I show that an optimal implementation of Q^1 leaves no room for joint-manipulations by high-skilled and low-skilled individuals. The argument for Q^2 would be the same. An optimal implementation of Q^1 implies that, for all t,

$$V_t^0 := V_t(s_{LL}) = V_t(s_{LH}) \quad \text{and} \quad V_t^1 := V_t(s_{HL}) = V_t(s_{HH}) \ .$$

Now suppose that the true state is s_{LL} and that high-skilled and low-skilled would prefer to induce jointly the outcome foreseen for state s_{HH}. However, the low-skilled would be able to induce this outcome without the help of high-skilled individuals and choose not to do so because of (4). Hence, this collective deviation does not satisfy the unanimity requirement. The same reasoning applies if the true state is s_{LH}, s_{HL} or s_{HH}.

Finally, there are the provision rules Q^* defined by the property $Q(s) = 0 \Longleftrightarrow s = s_{LL}$ and the provision rule Q' defined by $Q(s) = 1 \Longleftrightarrow s = s_{LL}$. I show that under provision rule Q^* high-skilled and low-skilled individuals have no incentive to deviate jointly. Propositions 5 and 6 (see section 5) imply that if an implementation of Q^* is optimal, then for all t,

$$V_t^0 := V_t(s_{LL}) \quad \text{and} \quad V_t^1 := V_t(s_{HL}) = V_t(s_{LH}) = V_t(s_{HH}) \ .$$

But this implies that the argument given for provision rule Q^1 is applicable; i.e. the *C-RT* conditions in (4) and (5) imply that joint deviations by high and low-skilled individuals are not attractive. ■

[19] Since it entails no loss of generality to assume that any interest group reveals the skill levels of its members, this condition is both necessary and sufficient for the elimination of any interest group that contains high-skilled and low skilled individuals who jointly induce $s_{\hat{x}\hat{y}}$, see the proof of Proposition 2.

Proof of Proposition 4

By plugging provision rule Q^* into (4) and (5), one arrives at the following set of conditions; $V_1(s_{LH}) = V_1(s_{HH})$, $\theta_H \geq V_1(s_{LL}) - V_1(s_{HL}) \geq \theta_L$, $V_2(s_{HL}) = V_2(s_{HH})$ and $\theta_H \geq V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L$. The optimal admissible social choice function satisfies (4) and (5) only if these statements remain true as one replaces $V_t(s)$ by $U_t(0)$ if $s = s_{LL}$ and by $U_t(k)$ if $s \neq s_{LL}$; that is, if and only if $\theta_H \geq U_1(0) - U_1(k) \geq \theta_L$ and $\theta_H \geq U_2(0) - U_2(k) \geq \theta_L$. This is definition of *Scenario 1.*

■

Proof of Lemma 2

Consider the Lagrangean of problem (6):

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

Deriving first order conditions, one easily verifies that *(BC)* has to be binding, that there is *no distortion at the top* and that there is a *distortion at the bottom* if and and only if (RS_2) is binding.

Denote by $(\bar{Y}_1(\bar{V}_2), \bar{C}_1(\bar{V}_2), \bar{Y}_2(\bar{V}_2), \bar{C}_2(\bar{V}_2))$ the solution of optimization problem (6). The 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 functions of $\bar{V}_2$. The fact that *(BC)* is binding yields a unique iso-tax-revenue line $Y_1 - C_1 = \gamma$, with $\gamma = C_2(\bar{V}_2) - Y_2(\bar{V}_2) + r$. $(\bar{Y}_1(\bar{V}_2), \bar{C}_1(\bar{V}_2))$ 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, this problem has a unique solution.

Denote the optimal values of the multipliers at the solution of problem (6) by $\bar{\lambda}(\bar{V}_2)$ and $\bar{\nu}(\bar{V}_2)$. The first order conditions imply:

$$\bar{\lambda}(\bar{V}_2) = \frac{1 - \overline{MRS}_1}{1 - \widehat{MRS}} \quad \text{and} \quad \bar{\nu}(\bar{V}_2) = \frac{u'(\bar{C}_1)}{u'(\bar{C}_2)} \frac{\overline{MRS}_1 - \widehat{MRS}}{1 - \widehat{MRS}} \,, \tag{15}$$

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

$\bar{\lambda}(\bar{V}_2) \geq 0$ captures the effect that a lower level of $\bar{V}_2$ tends to reduce V_1^* because of 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 V_1^* as less resources are needed to equip type 2 individuals with a utility level of $\bar{V}_2$.

These multipliers are used to study how V_1^* depends on $\bar{V}_2$. The following property is used:

$$\frac{\partial V_1^*}{\partial \bar{V}_2} = \bar{\lambda}(\bar{V}_2) - \bar{\nu}(\bar{V}_2) \,.$$

Similarly as in the proof of Lemma 1, comparative statics of the solution of problem (6) with respect to $\bar{V}_2$ can be derived. Based on this exercise, the comparative statics of the Lagrangean multipliers can be determined. The details of this comparative statics

exercise are left to the reader. One arrives at the following results:
Step 1. Suppose first that (RS_2) is binding. Using the premise of Proposition 1, one verifies that the function $\bar{\lambda}(\bar{V}_2)$ decreases in $\bar{V}_2$ and that the function $\bar{\nu}(\bar{V}_2)$ increases in $\bar{V}_2$, i.e. as long as (RS_2) is binding the function V_1^* is strictly concave in $\bar{V}_2$ and one has:

$$\frac{\partial^2 V_1^*}{\partial(\bar{V}_2)^2} \quad = \quad \bar{\lambda}'(\bar{V}_2) - \bar{\nu}'(\bar{V}_2) \quad < \quad 0 \, .$$

Step 2. Assume that (RS_2) is not binding.[20] The first order conditions imply $\bar{\lambda}(\bar{V}_2) = 0$ and $\bar{\nu}(\bar{V}_2) = u'(\bar{C}_1)/u'(\bar{C}_2)$. The comparative statics with respect to $\bar{V}_2$ reveal

$$\frac{\partial^2 V_1^*}{\partial(\bar{V}_2)^2} \quad = \quad -\bar{\nu}'(\bar{V}_2) \quad < \quad 0 \, .$$

Step 3. As $\bar{\lambda}(\bar{V}_2)$ decreases in $\bar{V}_2$, there is a critical value $\hat{U}_2$, such that if this critical value is exceeded, (RS_2) is no longer binding. Moreover, one can show that the function $\bar{\lambda}(\bar{V}_2) - \bar{\nu}(\bar{V}_2)$ is continuous at $\hat{U}_2$.
Step 4. V_1^* has a maximum. This follows from the existence of a solution of the following problem:

$$\begin{array}{ll} \max_{C_1,Y_1,C_2,Y_2} & u(C_1) - v\left(\frac{Y_1}{w_1}\right) \\ \text{s.t.} & Y_1 - C_1 + Y_2 - C_2 \geq k \, , \\ & u(C_2) - v\left(\frac{Y_2}{w_2}\right) \geq u(C_1) - v\left(\frac{Y_1}{w_2}\right) \, . \end{array} \tag{16}$$

Denote by $\tilde{V}_2$ the utility level that results for type 2 individuals at a solution to problem (16). Using the first order conditions of problem (16) allows us to verify that $\bar{\lambda}(\tilde{V}_2) - \bar{\nu}(\tilde{V}_2) = 0$.
Step 5. Using optimality condition (10) for the *optimal admissible allocation* to substitute for $u'(\bar{C}_1)/u'(\bar{C}_2)$ in the formula for $\bar{\nu}(\bar{V}_2)$ (see (15)), one gets $\bar{\lambda}(U_2(r)) - \bar{\nu}(U_2(r)) = -1$.

■

Proof of Proposition 5

By plugging provision rule Q^* into (4) and (5), one arrives at the following set of conditions; $V_1(s_{LH}) = V_1(s_{HH})$, $\theta_H \geq V_1(s_{LL}) - V_1(s_{HL}) \geq \theta_L$, $V_2(s_{HL}) = V_2(s_{HH})$ and $\theta_H \geq V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L$.
I consider a relaxed version of the problem to maximizing EW subject to admissibility and (4) and (5). The relaxed problem takes only the following collective incentive constraints into account: $V_1(s_{LH}) = V_1(s_{HH})$, $V_2(s_{HL}) = V_2(s_{HH})$ and $V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L$. I will show below that the solution to this relaxed problem is also a solution of the "full" problem. I denote by $\{V_{t,x}^{**}(s)\}_{s \in S}$ the utility levels realized by individuals with skill level w_t at a solution to the relaxed problem.

[20] The existence of a value of $\bar{V}_2$ such that (RS_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 $Y_t = C_t + \frac{k}{2}$. The resulting allocation is efficient by the first welfare theorem.

Assumption 3 $\theta_L < \bar{\theta} := V_1^*(U_2(k) + \theta_L, 0) - V_1^*(U_2(0) - \theta_L, r)$.

The assumption ensures that a solution of the relaxed problem does not violate the neglected collective incenitve constraint for less productive individuals.[21]

Step 1. For every s, the budget constraint is binding and there is *no distortion at the top*. This follows from setting up the Lagrangean and deriving first order conditions.

Step 2. Verify that $V_{2,x}^{**}(s_{LL}) \leq \theta_L + U_2(k)$ and that $V_{2,x}^{**}(s_{LH}) \leq U_2(k)$: Suppose, that $V_{2,x}^{**}(s_{LL}) > \theta_L + U_2(k)$. Suppose the planner would instead choose $(Y_1^*(k), C_1^*(k)$ and $Y_2^*(k), C_2^*(k))$ for $s \neq s_{LL}$. For $s = s_{LL}$, the planner would choose $V_{2,x}^{**}(s_{LL}) = \theta_L + U_2(k)$ and $V_1 = V_1^*(\theta_L + U_2(k), 0)$. Due to the monotonicity properties established in Lemma 2, this would increase utilitarian welfare in every state s. The constraints in (5) imply $V_{2,x}^{**}(s_{LH}) \leq V_{2,x}^{**}(s_{LL}) - \theta_L$. Hence, $V_{2,x}^{**}(s_{LH}) \leq U_2(k)$.

Step 3. Verify that $V_{1,x}^{**}(s_{LL}) = V_1^*(V_{2,x}^{**}(s_{LL}), 0)$ and $V_{1,x}^{**}(s_{HL}) = V_1^*(V_{2,x}^{**}(s_{HL}), k)$; i.e. these allocations lie on the Pareto frontier: Note that $V_{1,x}^{**}(s_{LL}) \neq V_1^*(V_{2,x}^{**}(s_{LL}), 0)$ or $V_{1,x}^{**}(s_{HL}) \neq V_1^*(V_{2,x}^{**}(s_{HL}), k)$ would immediately yield a contradiction to optimality.

Step 4. Verify that $V_{2,x}^{**}(s_{LL}) \geq U_2(0)$ and $V_{1,x}^{**}(s_{LL}) \leq U_1(0)$: If this is false, then $V_{2,x}^{**}(s_{LL}) < U_2(0)$ and $V_{1,x}^{**}(s_{LL}) > U_1(0)$ by *Step 3*. Then, using the monotonicity properties established in Lemma 2, it is possible to increase $V_2(s_{LL})$ and to decrease $V_1(s_{LL})$ along the 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}.

Step 5. Verify that $V_{2,x}^{**}(s_{LH}) = V_{2,x}^{**}(s_{HH}) = V_{2,x}^{**}(s_{HL}) =: V_{2,x}^{**}(s_H) \leq U_2(k)$: $V_{2,x}^{**}(s_{HH}) = V_{2,x}^{**}(s_{HL})$ follows from (5), and $V_{2,x}^{**}(s_{LH}) \leq U_2(k)$ has been established in *Step 2*. 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 that the utility levels $V_{1,x}^{**}(s_{LH}) = V_{1,x}^{**}(s_{HH})$ are equal to those that are 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 admissibility constraints

$$\begin{aligned} &Y_1(s_{LH}) - C_1(s_{LH}) + Y_2(s_{LH}) - C_2(s_{LH}) \geq r \quad (\text{BC}(s_{LH}))\,, \\ &Y_1(s_{HH}) - C_1(s_{HH}) + Y_2(s_{HH}) - C_2(s_{HH}) \geq r \quad (\text{BC}(s_{HH}))\,, \end{aligned}$$

and

$$\begin{aligned} &u(C_1(s_{LH})) - v\left(\frac{Y_1(s_{LH})}{w_2}\right) \leq V_{2,x}^{**}(s_{LH}) \quad (\text{RS}_2(s_{LH}))\,, \\ &u(C_1(s_{HH})) - v\left(\frac{Y_1(s_{HH})}{w_2}\right) \leq V_{2,x}^{**}(s_{HH}) \quad (\text{RS}_2(s_{HH}))\,; \end{aligned}$$

[21] Recall that under *Scenario 2*, $\theta_L > U_2(0) - U_2(k)$. If θ_L does not exceed $U_2(0) - U_2(k)$ by too much, i.e. $\theta_L \simeq U_2(0) - U_2(k)$, then, the assumption $\theta_L < V_1^*(U_2(k) + \theta_L, 0) - V_1^*(U_2(0) - \theta_L, k)$ is satisfied. To see this: if $\theta_L \simeq U_2(0) - U_2(k)$, the continuity property established in Lemma 2 implies that $V_1^*(U_2(k) + \theta_L, 0) - V_1^*(U_2(0) - \theta_L, k) \simeq U_1(0) - U_1(k)$. By definition of *Scenario 2*, the latter term exceeds θ_L.

and the collective incentive 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, finally, 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 proceed as follows to arrive at a contradiction. (i) 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. (ii) 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 (BC(s_{HH})) and (RS$_2$(s_{LH})). (iii) Use the geometry of this set, the strict quasiconcavity of preferences, as well as the fact that $V^{**}_{2,x}(s_{LH}) \leq U_2(k)$, established in *Step 2*, 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 (BC(s_{HH})) and (RS$_2$(s_{LH})) are binding, while (BC(s_{LH})) and (RS$_2$(s_{HH})) are slack. Finally, note that a strict inequality in (BC(s_{LH})) contradicts *Step 1.*

Step 6. Note that *Steps 4* and 5, the monotonicity properties established in Lemma 2 and $V^{**}_{2,x}(s_H) \leq U_2(k)$ imply that $V^{**}_{1,x}(s_{LH}) = V^{**}_{1,x}(s_{HH}) = V^{**}_{1,x}(s_{HL}) =: V^{**}_{1,x}(s_H) = V^*_1(V^{**}_{2,x}(s_H), k) \geq U_1(k)$. Moreover, at a solution to the relaxed problem, $(C_1(s_{HL}), Y_1(s_{HL})) = (C_1(s_{LH}), Y_1(s_{LH})) = (C_1(s_{HH}), Y_1(s_{HH})) =: (C_1(s_H), Y_1(s_H))$ and $(C_2(s_{HL}), Y_1(s_{HL})) = (C_2(s_{LH}), Y_1(s_{LH})) = (C_2(s_{HH}), Y_1(s_{HH})) =: (C_2(s_H), Y_2(s_H))$ due to the uniqueness established in Lemma 2.

Step 7. Verify that $V^{**}_{1,x}(s_H) \neq U_1(k)$ and $V^{**}_{2,x}(s_H) \neq U_2(k)$ and $V^{**}_{2,x}(s_{LL}) \neq U_2(0)$ and $V^{**}_{1,x}(s_{LL}) \neq U_1(0)$ and $V^{**}_{2,x}(s_{LL}) - V^{**}_{2,x}(s_{LH}) = \theta_L$.

This follows from setting up the Lagrangean 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 optimal admissible allocation, 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.

Step 8. Verify that, at solution to the relaxed problem, the neglected collective incentive constraint for low-skilled individuals is satisfied; i.e. $\theta_H > V^{**}_{1,x}(s_{LL}) - V^{**}_{1,x}(s_H) > \theta_L$

First, note that by *Steps* 4,6,7 and the definition of *Scenario 2*, $\theta_H > U_1(0) - U_1(k) > V^{**}_{1,x}(s_{LL}) - V^{**}_{1,x}(s_H)$. Second, note that $V^{**}_{1,x}(s_{LL}) \geq V^*_1(U_2(k) + \theta_L, 0)$ and $V^{**}_{1,x}(s_H) \leq V^*_1(U_2(0) - \theta_L, r)$. The first inequality follows from *Steps* 2 and 3 and the monotonicity property established in Lemma 2. The second inequality is established as follows: Analogously to *Step* 2, one shows that $V^{**}_{2,x}(s_H) \geq U_2(0) - \theta_L$ and then uses $V^{**}_{2,x}(s_H) = V^*_1(V^{**}_{2,x}(s_H), k)$ and again the monotonicity property. Finally, using Assumption (3) establishes the result. ■