On the legitimacy of coercion for the financing of public goods*

Felix Bierbrauer[†]

University of Cologne

December 5, 2011

Abstract

The theory of public-goods provision under conditions of incomplete information has established an impossibility result: One cannot have efficient outcomes and voluntary participation at the same time. This paper asks whether priority should be given to outcomes that are efficient or to outcomes that avoid the use of coercion. The main result is that the answer depends on the quality of institutions. If there are no effective barriers to rent extraction by policy-makers, it is desirable that individuals are protected against an abuse of coercive power, even if this implies that public-goods provision is inefficient.

Keywords: Public Goods, Political Economy, Mechanism Design, Participation Constraints

JEL: D02, D82, H41

---

*I benefited from conversations with Pierre Boyer, Alia Gizatulina, Martin Hellwig, Jos Jansen, Philipp Weinschenk, and John Weymark. Moreover, I am grateful for the comments of two anonymous referees and an editor.

[†]University of Cologne, Center for Macroeconomic Research, Albert-Magnus- Platz, 50923 Cologne, Germany. Email: bierbrauer@wiso.uni-koeln.de

1 Introduction

Should public goods be financed solely out of voluntary contributions or is there a role for taxes as a source of public goods finance? The disadvantage of a system based on voluntary contributions is that public goods are underprovided because voluntary contributions tend to neglect the social benefits of increased public-goods provision. By contrast, a system based on taxes which are raised independently of an individual's valuation of public goods can provide sufficient funds for efficient public-goods provision. However, if individuals can be forced to pay for public goods that they do not value, this coercive power may also be abused. If politicians, bureaucrats or managers can use funds from the public budget to finance projects, their choices may be biased towards their private interests.

This paper formalizes a tradeoff between the efficiency of public-goods provision, on the one hand, and protection against an abuse of coercive power, on the other. It asks the question under what circumstances coercion is legitimate in the sense that it makes individuals, by and large, better off even though occasionally they will be forced to pay for a public good that they do not value.

The analysis is based on a model of mechanism design; that is, we avoid any a priori assumption on the set of admissible policies. However, and in contrast to most of the literature that applies tools of mechanism design to the study of public-goods provision, we will assume that the institution in charge of organizing the supply of public goods, henceforth simply referred to as the *mechanism designer*, is, at least partly, self-interested and tries to channel tax revenues into his own pockets. Our answer to the question whether coercion is legitimate will depend on a measure of institutional quality that affects the fraction of tax revenues that the mechanism designer will be able to keep for himself.

The logic of the argument is as follows: Any mechanism designer will provide a surplus-maximizing amount of public goods if coercion is possible. If the quality of institutions is low so that the designer can act as a malevolent Leviathan, he will keep the entire surplus for himself and individuals will not benefit from public-goods provision at all. Voluntary public goods finance, by contrast, leaves them at least a positive share of the surplus. Individuals prefer a positive share of a small surplus over a negligible share of a large surplus. Consequently, coercion is legitimate if and only if the mechanism designer is prevented from extracting the surplus from public-goods provision; i.e. if and only if the institutional quality is sufficiently high.

What does this mean for applications? To see what is at stake when we talk about efficient versus second-best public-goods provision it is instructive to look at specific applications.

People who want their houses to be protected may privately hire security services, or the state may send police-men. State provision may in principle be preferable for a variety of reasons: There may be scale effects in the provision of security to many neighborhoods, or the citizens of a country agree that the protection against criminal activity should be a universal service provided to all, and not just to those who can afford it etc. If, however, the coercive power of the police is used primarily for the purpose of rent-extraction, e.g. if the police is corrupt or even captured by criminal organizations, then not having a police at all, may be the preferable outcome.

If individuals are altruistic in the sense that they care for the well-being of the needy, then charitable donations can be thought of as voluntary contributions to a public good. However, the typical prediction of an economic theorist would be that the outcome that is thereby induced is inefficient because donors will not internalize the full social benefit of their donation.[1] An efficient outcome may therefore require to have a tax-financed welfare state.[2] This argument relies on the assumption that the welfare state would indeed realize the efficiency gains that are left over by a system based on voluntary contributions. However, if the functioning of the welfare state is impaired by political economy forces which imply that resources are channeled to the middle-class rather than to the needy, the conclusion that state-provision of welfare dominates private provision requires a more elaborate analysis that explicitly takes these political economy forces into account.

Goods and services that are non-rival but excludable, such as streets, telecommunications networks, or TV programs provide a further class of applications. First-best efficiency requires that the possibility to exclude individuals from the use of such a good should never be used. It is inefficient if people have to pay a toll if they want to cross a bridge. Some people will avoid the bridge, even though admitting them would be costless.

[1] See e.g. Bergstrom et al. (1986).

[2] The prediction of Bergstrom et al. (1986) would be that the introduction of such an efficient welfare state implies a complete crowding out of voluntary contributions to charity.

However, if the bridge is not financed out of user fees, the state has to extract tax revenue somewhere else. Now, if these tax revenues are used for all sorts of purposes other than the financing of infrastructure, or if the location of bridges is biased towards the districts of particular politicians, then it may be preferable to have a private provision of bridges which relies on user-fees. This may not be first-best, but it may ensure that bridges are build at places where user-fees can be generated; that is, at places where a sufficient number of people demand the services provided by a bridge.

A sketch of the formal analysis. The literature on public-goods provision under conditions of incomplete information about public goods preferences has arrived at two major results. On the one hand, there is a *possibility result*: It is possible to reach an efficient allocation of public goods, even if individuals have private information on their preferences.[3] On the other hand, there is an *impossibility result*: Efficient outcomes are out of reach if participation constraints have to be respected, so that each individual has to be better off relative to a status quo allocation without public-goods provision.[4]

This literature confronts us with a choice between outcomes that are efficient and second-best outcomes that avoid the use of coercion. However, it does not discuss how the choice should be made. A naive reading of this literature leads to the conclusion that coercion is always desirable because otherwise efficiency cannot be reached. The contribution of the present paper is to enrich this framework by looking more carefully at the objective of the mechanism designer and to show that the use of coercion is not generally desirable.

We take a constitutional choice perspective to see whether the objective to reach efficient outcomes justifies the use of coercion; that is, we make the following thought experiment: Suppose there is an ex ante stage at which individuals have not yet discovered what their preferences are. More specifically, an individual's objective is to maximize expected utility, with expectations taken about her future preferences. At this ex ante stage, individuals decide about the rules according to which public goods are provided.

[3]This result is due to d'Aspremont and Gérard-Varet (1979). For a more recent generalization, see d'Aspremont et al. (2004).

[4]See Güth and Hellwig (1986), Rob (1989), Mailath and Postlewaite (1990), Hellwig (2003), or Neeman (2004).

In particular, they face a choice between a strong and a weak formulation of participation constraints. The strong formulation requires that, at the interim stage, where individuals have discovered their preferences, each individual benefits from public-goods provision. The weak version requires only that individuals benefit from public-goods provision at the ex ante stage.

To illustrate this by means of an example, think of the construction of a bridge, and suppose that there are individuals who cross the bridge frequently and others who do so only rarely. If we impose participation constraints in the weak, ex ante sense, this allows us to force the non-frequent users to contribute to the financing of the bridge, provided that their utility loss is compensated for by the utility gain of the frequent users. By contrast, if we impose participation constraints in the strong, ex interim sense, we lose this opportunity. In this case, the less frequent users must also be made better off by the construction of the bridge, which implies that they cannot be forced to pay for a bridge that they hardly ever use.

We say that coercion is legitimate if, at the constitutional stage, individuals opt for participation constraints in the weak sense. The main insight of the paper is that strong participation constraints, which protect individuals from having to pay for a public good that they do not value, should be imposed if and only if there is a pronounced agency conflict between individuals and the mechanism designer.[5]

The crucial step in the formal argument is the observation that the difference between the payoffs that individuals realize with strong and weak participation constraints, respectively, depends on the mechanism designer's degree of benevolence. With weak participation constraints public goods are provided in such a way that a first-best surplus maximum is achieved. With strong participation constraints, by contrast, the realized surplus is only second-best. Despite their detrimental consequences from an efficiency perspective, strong participation constraints ensure that individuals get at least an information rent if public goods are provided. If the mechanism designer can extract the whole surplus from public-goods provision for himself, then the imposition of strong par-

[5]Hellwig (2003) studies public-goods provision by a benevolent mechanism designer who faces participation constraints in the strong sense. Our result shows that such a mechanism design problem cannot occur if the relevant constraints are endogenized by means of a constitutional decision at the ex ante stage.

ticipation constraints is desirable because individuals prefer getting an information rent over not getting anything.

The remainder of the paper is organized as follows. The next section gives a more detailed literature review. Section 3 introduces the model. As a benchmark, Section 4 shows that a mechanism design analysis gives rise to the conclusion that efficiency and voluntariness of participation are incompatible. Section 5 establishes the main result that the legitimacy of coercion depends on the mechanism designer's degree of benevolence. The last section contains concluding remarks. All proofs are in the Appendix.

2 Related Literature

This section discusses related literature. It has two parts. The first part provides a summary of the main results that the mechanism design literature has provided, to the extent that they are relevant for study of allocation problems in the public sector. In the second part we connect the present paper both with classical and recent work in the field of political economy.

Mechanism Design. There are two classical contributions to the theory of public goods which both claim that it is possible to provide public goods in an efficient and voluntary way; that is, efficient outcomes can be achieved without recourse to the state's coercive power.

On the one hand, there is Lindahl's (1919) proposal of a voluntary exchange theory of public finance. Lindahl assumes that, when confronted with appropriately calibrated personalized prices, individuals will demand that public goods are provided in an efficient way. This construction ensures, in particular, that all individuals prefer the outcome of a "Lindahl equilibrium" over a status quo situation with no public goods provision. We can therefore think of this equilibrium as resulting from a voluntary exchange. On the other hand there is the Coasian (1960) idea, that, whenever the provision of a public good is efficient, then individuals will find some bargaining procedure that makes it possible to realize these efficiency gains. The reliance on bargaining among the beneficiaries of public-goods-provision implies, once more, that coercive power will not be needed.

The theory of mechanism design provides a conceptual framework which makes it possible to study these ideas in a formal way. Basically, mechanism design theory has shown that, under the assumption that individuals are privately informed about their valuations of public goods, Lindahl and Coase are both wrong.[6] The results of Myerson and Satterthwaite (1983), Güth and Hellwig (1986), Mailath and Postlewaite (1990) and Hellwig (2003) provide conditions under which there exists no mechanism which yields efficient outcomes, is incentive-compatible thereby respecting the individual's private information on their preferences, and moreover, is such that participation is voluntary in the sense that no individual is made worse off by public-goods provision. This impossibility result shows the strength of the theory of mechanism design: It provides conclusions which apply to any conceivable bargaining protocol. Whatever the bargaining procedure, if participation is voluntary, then efficiency is out of reach.

By contrast, if it is possible to force people to pay for public goods, that they may not like, then efficiency can be reached, even if individuals are privately informed about their preferences. This result is due to d'Aspremont and Gérard-Varet (1979) and Arrow (1979). For a public finance economist, this result provides a possible justification for state interventions: Coercive power is needed in order to provide public goods efficiently. This argument, however, neglects the political economy forces that come into play whenever the state is doing something.

Political Economy. This paper follows Buchanan and Tullock (1962) in that it provides a formalization of the question which rules should govern the allocation of resources in the public sector. In contrast to Buchanan and Tullock, who analyze optimal majority rules, we do not focus on conflicts of interest between individuals with different preferences, but on the distributive conflict between politicians who try to extract rents from holding office, on the one hand, and the individuals who wish to consume public goods, on the other. To spell out this distributive conflict most clearly, we follow Brennan and Buchanan (1980) and model political institutions as a revenue-maximizing Leviathan.

We do, however, assume that institutional barriers may limit the politicians' ability

[6] has argued that individuals would never reveal the information that is required in order to compute the Lindahl prices. Rather, individuals would understate their preferences so as to make sure that they bear a smaller part of the provision cost of public-goods provision. However, Samuelson did not provide a formal treatment of this concern.

to extract revenues. The literature has studied a variety of mechanisms that may serve this purpose. In the classical political agency model by Barro (1973) and Fehrejohn (1986), voters can discipline politicians by threatening not to reelect them should they run away with the economy's resources. In a federal system with several local policy-makers, voters can use relative performance evaluation to identify politicians who have enriched themselves, see Wrede (2001). Alternatively, policy makers may have internalized social norms so that extracting tax revenues for themselves comes with a psychic cost even if it remains undetected by the electorate, Hillman and Ursprung (2000).[7]

This paper is based on these ideas. However, we do not invoke a specific model to explain how the taming of Leviathan is achieved. Instead, we look at a reduced form and parameterize the policy-maker's ability to extract rents. The advantage of this approach is that we can provide a comparative statics analysis that shows how the ideal institutional setting varies with the intensity of the political agency problem.

Two recent papers have also linked mechanism design theory and political economy. Grüner and Koriyama (2011) ask the question whether an efficient allocation of public goods can be achieved if it is required that each individual is made better off relative to a status quo situation with majority voting about public-goods provision. In this model, efficient public-goods provision may be possible. This result is given a positive interpretation; i.e., it explains why despite the impossibility results in the mechanism design literature, public goods are provided in the real world and why these outcomes may even be efficient: In the real world the status quo outcome is shaped by democratic institutions. The present paper offers an alternative positive explanation. Individuals may be willing to accept that, occasionally, they have to pay for public goods that they do not value if, on average, they benefit from the provision of public goods. This leads to a weaker notion of participation constraints, so that efficient outcomes can be achieved.

Acemoglu et al. (2008) compare public and private provision of insurance contracts. They formalize the following tradeoff: private provision suffers from inefficiencies due to participation constraints. These problems may be overcome by state provision of insurance, given that the state has coercive power. The disadvantage of state provision, however, is that the coercive power may be abused by selfish politicians. The tradeoff

[7]See Persson and Tabellini (2000) for a broader overview of political agency models.

"markets versus governments" therefore becomes a tradeoff between distortions due to participation constraints and distortions due to agency problems between citizens and politicians. While this tradeoff seems empirically plausible, the present paper shows that, from a normative perspective, agency problems are what justifies the imposition of participation constraints. Hence, why should the state be given coercive power if politicians are not acting in the citizen's interest? The normative analysis in this paper suggests that the state should have coercive power only if the agency problems between citizens and politicians are less significant than the agency problems between private providers of insurance and their customers.

Finally, this paper draws on the empirical literature which relates measures of economic performance to the quality of political institutions. In our theoretical treatment, performance is the same as Pareto-efficiency and the quality of institutions is identified with an office-holder's ability to extract monetary rents.[8] Moreover, the theoretical model is based on the primitive assumption that better political institutions generate better economic outcomes. The empirical literature provides us with ample evidence of this relationship. Examples include Adam et al. (2011) and Efendic et al. (2011).

3 The environment

There is a finite set of individuals, $I = \{1, \ldots, n\}$. The preferences of individual i are given by the utility function

$$u_i = \theta_i q - t_i,$$

where $q \in \mathbb{R}_+$ is the provision level of a public good, t_i is individual i's contribution to the cost of public good provision and θ_i is a taste parameter that affects individual i's valuation of the public good. For each i, θ_i belongs to a finite ordered set $\Theta = \{\theta^0, \theta^1, \ldots, \theta^m\}$, with $\theta^0 = 0$. We assume that $\theta^l - \theta^{l-1} = 1$, for all l. We denote a vector of all individual taste parameters by $\theta = (\theta_1, \ldots, \theta_n)$.

From an ex ante perspective, the taste parameters of individuals are independent and identically distributed random variables that take values in Θ. For any i, we denote the probability that $\theta_i = \theta^l$ by p^l. The following notation will prove helpful. For every i, let

[8] See, for instance, Becker et al. (2009) for evidence on corruption.

$p(\theta_i)$ be a random variable that takes the value p^l if θ_i takes the value θ^l and $P(\theta_i)$ be a random variable that takes the value $\sum_{k=0}^{l} p^k$ if θ_i takes the value θ^l. Define $h^l = \frac{1-P(\theta^l)}{p(\theta^l)}$. In the literature this fraction is known as the hazard rate. We assume that the hazard rate is decreasing, $h^l < h^{l-1}$, for all $l \geq 1$. This assumption is imposed in the following without further mention.

We study public goods provision from an interim perspective, i.e., after individuals have learned what their preferences are. With an appeal to the Revelation Principle we limit attention to direct mechanisms and to truthful Bayes-Nash equilibria. A direct mechanism consists of a provision rule for the public good and, for each individual i, a payment rule. The provision rule is a function $q : \Theta^n \to \mathbb{R}_+$ that specifies a public good provision level as a function of the preferences that individuals communicate to the mechanism designer. Analogously, the payment rule for individual i is a function $t_i : \Theta^n \to \mathbb{R}$.

A mechanism has to satisfy participation constraints, incentive compatibility constraints and a budget constraint. The budget constraint requires that expected payments of individuals are sufficient to cover the expected cost of public good provision,[9]

$$E\left[\sum_{i=1}^{n} t_i(\theta)\right] \geq E[k(q(\theta))] \ . \tag{1}$$

where k is a strictly increasing and strictly convex function with $k(0) = 0$, $\lim_{q\to 0} k'(q) = 0$, and $\lim_{q\to\infty} k'(q) = \infty$. The expectations operator E applies to the vector θ of all individual taste parameters.

The incentive compatibility constraints ensure that that truth-telling is a Bayes-Nash equilibrium: given that all other individuals reveal their taste parameter, the best response of individual i is to reveal the own taste parameter as well. Formally, for each i, for each $\theta_i \in \Theta$, and for each $\hat{\theta}_i \in \Theta$,

$$\theta_i Q_i(\theta_i) - T_i(\theta_i) \geq \theta_i Q_i(\hat{\theta}_i) - T_i(\hat{\theta}_i), \tag{2}$$

where

$$Q_i(\hat{\theta}_i) := E[q(\theta_{-i}, \hat{\theta}_i) \mid \hat{\theta}_i]$$

[9] It has been shown by d'Aspremont et al. (2004) that to any incentive compatible mechanism that satisfies the budget constraint in expectation there exists a payoff equivalent incentive compatible mechanism that satisfies budget balance in an ex post sense, i.e., $\sum_{i=1}^{n} t_i(\theta) \geq k(q(\theta))$, for every θ. Hence, working with budget balance in expectation is without loss of generality.

is the expected level of public good provision from the perspective of individual i, given that all other individuals reveal their preferences to the mechanism designer and individual i announces $\hat{\theta}_i$. Likewise,

$$T_i(\hat{\theta}_i) = E[q(\theta_{-i}, \hat{\theta}_i) \mid \hat{\theta}_i]$$

is i's expected payment.

A mechanism also has to satisfy participation constraints which ensure that individuals benefit from the provision of the public good. We distinguish between participation constraints at the ex interim state or at the ex ante stage. The ex interim participation constraints are as follows: For all i, and all θ_i,

$$\theta_i Q_i(\theta_i) - T_i(\theta_i) \geq 0. \tag{3}$$

These constraints ensure that, after all individuals have discovered what their public goods preferences are, no individual is worse off relative to a status quo situation in which the public good is not provided. An alternative interpretation is that individuals are given veto rights that protect them from having to pay for a public good that they do not value. Consequently, a deviation from the status quo requires a unanimous agreement to provide the public good.

The ex ante participation constraints require that, for all i,

$$E[\theta_i q(\theta) - t_i(\theta)] \geq 0 \ , \tag{4}$$

so that each individual benefits from public good provision at an ex ante stage, i.e., prior to learning what the own preferences are. These participation constraints are less restrictive than those in (3). To make this more explicit we can use the law of iterated expectations to write (4) as follows: for all i,

$$\sum_{l=0}^{m} p^l (\theta^l Q_i(\theta^l) - T_i(\theta^l)) \geq 0 \ .$$

Consequently, the ex ante participation constraints in (4) require that the ex interim participation constraints in (3) hold "on average", but not necessarily for each possible realization of individual i's preferences.

These constraints in (4) ensure that the provision of the public good is a Pareto-improvement if considered behind a "veil of ignorance" where individuals can form an expectation about how the public good is going to affect their well-being, but are not yet fully informed about their preferences. They make it possible to rely on coercion when financing the provision of a public good. Individuals can be forced to pay for a public good that they do not value, provided that, behind the veil of ignorance, they benefit from public goods provision. The participation constraints in (3), by contrast, exclude coercion under each and every circumstance. We can therefore interpret them as providing a maximal protection of economic freedom: No one may interfere with an individual's decision to spend his money on the uses that are most attractive to him.

The analysis focusses on the question whether the use of coercion is beneficial for individuals. To this end we will compare mechanisms where the ex ante participation constraints have to be satisfied to mechanisms where the ex interim participation constraints are imposed. The standard of comparison is the ex ante expected utility of individuals. If this utility is larger with ex ante participation constraints, then we say that the use of coercion is legitimate in the sense that if, behind a veil of ignorance, individuals were confronted with a constitutional choice about the use of coercion, they would unanimously vote in favor of it.

4 The tradeoff between efficiency and voluntary participation

As a benchmark, we will show in the following that Pareto-efficiency is compatible with participation constraints at the ex ante stage but not with participation constraints at the ex interim stage. The analysis is institution-free in the sense that we only ask whether mechanisms with certain properties exist. We do not yet ask what mechanisms a policy-maker would propose, or whether individuals would be better off if they were given veto rights. This will be studied in the next section.

Our treatment generalizes the previous literature which has focused on surplus-maximizing outcomes. This literature has shown that the allocation that maximizes the social surplus or, equivalently, a utilitarian welfare function can no longer be reached if incentive compatibility constraints and participation constraints at the ex interim stage have to be respected. There are two problems with this: first, a utilitarian welfare or social surplus maximum is one among several Pareto-efficient allocations. The impossibility to reach a surplus maximum does not yet prove that it is impossible to reach any Pareto-efficient outcome. Second, there is the question whether the definition of Pareto-efficiency should invoke the individuals' ex ante or ex interim expected utility. If the move from a status quo without public goods to a surplus-maximizing mechanism is impossible without violating some interim participation constraints, this simply means that – from an ex interim perspective – the move is not a Pareto-improvement. Hence, with an ex interim notion of Pareto-efficiency, we cannot conclude that there is a tension between efficiency and voluntary participation.

We therefore invoke a notion of Pareto-efficiency that is based on expected utility from the ex ante perspective.[10] More formally, we say that a mechanism $(q, t_1, \ldots, t_n)$ is *Pareto-efficient* if it is incentive compatible and budgetary feasible, and there is no other incentive compatible and budgetary feasible mechanism $(q', t'_1, \ldots, t'_n)$, such that for all i, $E[\theta_i q'(\theta) - t'_i(\theta)] \geq E[\theta_i q(\theta) - t_i(\theta)]$, with a strict inequality for some i.

Proposition 1 *A mechanism is Pareto-efficient if and only if the budget condition (1) holds as an equality and the public-goods provision rule is surplus-maximizing; i.e., for every θ, $q(\theta)$ is chosen so as to maximize $(\sum_{i=1}^{n} \theta_i)\, q(\theta) - k(q(\theta))$.*

It is well known that, under conditions of complete information, surplus maximization in conjunction with budget balance is both necessary and sufficient for Pareto-efficiency if preferences are quasilinear in money. Proposition 1 shows that the same is true with private information on public-goods preferences, i.e., private information on preferences does not alter the efficiency conditions. Moreover, it provides a justification for the objective of surplus-maximization, which has been the focus of the previous literature.

It is now straightforward to show that Pareto-efficiency is not in conflict with ex ante participation constraints. The if-part of Proposition 1 implies that there exists a Pareto-efficient allocation such that the surplus from public-goods provision is shared equally among individuals, i.e., such that the ex ante expected utility of each individual i is equal

[10] For an ex interim notion of Pareto- efficiency, see Holmstrom and Myerson (1983); see Ledyard and Palfrey (1999) for an application to public-goods provision.

to

$$E[\theta_i q(\theta) - t_i(\theta)] = \frac{1}{n} E\left[\sum_{i=1}^{n} \theta_i q^*(\theta) - k(q^*(\theta))\right] > 0 \ ,$$

where q^* is the surplus-maximizing provision rule. In the following we will refer to this mechanism as the *symmetric efficient mechanism.* Obviously, since the expected surplus from public-good provision is strictly positive, under the symmetric Pareto-efficient mechanism all ex ante participation constraints hold as a strict inequality. Hence, if the expected benefits from public goods provision are evenly distributed, then every individual is made better off by public good provision.[11] We summarize these observations in the following Corollary.

Corollary 1 *There exists a Pareto-efficient mechanism that satisfies the ex ante participation constraints.*

By contrast, efficiency may be out of reach with ex interim participation constraints.

Proposition 2 *There exists a Pareto-efficient mechanism that satisfies the interim participation constraints if and only if*

$$E\left[\sum_{i=1}^{n} \left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q^*(\theta)\right] \geq E[k(q^*(\theta))] \ . \tag{5}$$

The term $\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}$ is known as individual i's virtual valuation of the public good. The virtual valuation is interpreted as the maximal amount that type θ_i of individual i will pay for the public good in the presence of incentive compatibility and ex interim participation constraints. In a complete information environment, the maximal amount would be the true valuation θ_i. With private information on public-goods preferences this expression is reduced by the hazard rate $\frac{1-P(\theta_i)}{p(\theta_i)}$. The hazard rate can therefore also be interpreted as an information rent. The Proposition says that Pareto-efficiency is possible if and only if, with a surplus-maximizing provision rule, the expected virtual valuation exceeds the expected resource costs of public goods-provision.

[11] There also exist Pareto-efficient mechanisms that violate the ex ante participation constraints. Given that utility is perfectly transferrable between individuals, efficiency can also be achieved if one individual is made very badly-off and receives a negative expected payoff.

It can be shown that the inequality in (5) is violated if the number of individuals is sufficiently large.[12] The intuition is that, as the number of individual's increases, each single individual's influence on the public goods provision level becomes smaller and smaller, so that it becomes more and more attractive to articulate a low taste parameter in order to mitigate the own contribution to the cost of public good provision. Hence, with more individuals it is more difficult to raise enough money for efficient public goods provision.[13]

The observation that ex interim participation constraints may imply that efficiency cannot be reached has led various authors to study second best mechanisms. Typically, the objective is to maximize expected utilitarian welfare, $E\left[\sum_{i=1}^{n}(\theta_i q(\theta) - t_i(\theta))\right]$, subject to the budget constraint (1), the incentive compatibility constraints (2), and the ex interim participation constraints (3). For brevity, we refer to this problem in the following as the *benevolent second-best problem.*

Proposition 3 *If condition (5) holds, then the symmetric Pareto-efficient mechanism solves the benevolent second-best problem. If condition (5) is violated, then the symmetric Pareto-efficient mechanism Pareto-dominates the solution to the benevolent second-best problem.*

The Proposition shows that from a normative perspective, the imposition of ex interim participation constraints is undesirable. These constraints are never beneficial but sometimes harmful, depending on whether or not the inequality in (5) holds. Hence, if individuals were given a choice between the imposition of participation constraints at the ex interim or the ex ante stage they would unanimously opt for the latter. Put differently, at the ex ante stage, individuals are happy to accept that they occasionally will have to pay for public goods that they do not value if they are assured that their share of the expected surplus from public goods provision is sufficiently high.

However, these results rely on the assumption that the mechanism designer is benevolent. Clearly, if the mechanism designer is such a nice guy, individuals have no reason

[12] A proof of this claim can be found in the supplementary material for this paper.

[13] For a model so that θ_i is distributed according to an atomless probability distribution, this impossibility result holds irrespective of the number of individuals, see Hellwig (2003).

to make his life harder by imposing ex interim participation constraints. To make sense of this requirement we therefore have to relax the assumption of benevolence.

5 Conflicts of interest

In the following, we allow for the possibility that the mechanism designer does not only care about the expected payoffs of individuals, but also derives utility from resources that he extracts for himself. In particular, we will formulate a model in which we can vary the mechanism designer's degree of benevolence. With this model we will ultimately show that coercion is legitimate if and only if the mechanism designer is sufficiently benevolent.

Formally, we assume that the public good is provided by a mechanism designer whose objective is to maximize net tax revenue,

$$R := E\left[\sum_{i=1}^{n} t_i(\theta)\right] - E[k(q(\theta))] \ , \tag{6}$$

i.e. the money that is left after the cost of public-goods provision is covered. A fraction $1 - \rho$ of these tax revenues are redistributed to individuals in a lump sum fashion. The politician keeps a fraction ρ for himself; that is, the monetary rent extracted by the politician equals

$$\rho\left(E\left[\sum_{i=1}^{n} t_i(\theta)\right] - E[k(q(\theta))]\right)$$

We treat ρ as a given parameter which measures the mechanism designer's degree of benevolence. High values of ρ indicate that the mechanism designer can pocket a large fraction of the tax revenues. In particular, we can interpret $\rho = 1$ as the case of a malevolent Leviathan who only cares for himself and $\rho = 0$ as indicating that the mechanism designer is fully benevolent. For ease of exposition, in case that the tax revenues exceed the cost of public-goods provision all individuals receive the same transfer. The per capita share of net tax revenue is therefore equal to $\frac{1-\rho}{n}$.

The parameter ρ measures the intensity of the agency conflict between the policy-maker and the individuals. This parameter could be endogenized by a more microfounded model. For instance, in the political agency model of Barro (1973) and Fehrejohn (1986) a ruling politician extracts a rent that is proportional to the economy's output. The fraction that he extracts depends on his discount factor since he is trading-off the rents he will enjoy in future periods should he stay in office and the rents that he could extract immediately, at the cost of being replaced at the next election. Other forces that limit a politician's ability or willingness to "steal" tax revenue include social norms, media attention, the intensity of political competition etc. However, for the purposes of this paper, modeling these forces explicitly would lead us astray. We seek to show that the intensity of the agency conflict determines whether the use of coercion is legitimate. This argument does not depend on the details of a specific political game that could be used to explain how the agency conflict arises.

The interaction between the mechanism and the individuals who consume the public good proceeds as follows. First, prior to the operation of the mechanism, the mechanism designer makes an unconditional lump sum payment of $\frac{1-\rho}{n}R^*$ to each individual, where R^* are the expected tax revenues that result from the mechanism design problem. After this payment is made, a mechanism is chosen in order to maximize R subject to the incentive compatibility constraints in (2), and participation constraints. Again, the participation constraints are either imposed at the ex ante stage as in (4), or at the ex interim stage as in (3).

5.1 A discussion of alternative modeling choices

An alternative sequential structure. Imposing a sequential structure where expected tax revenues are distributed prior to the operation of the mechanism has the following convenient implication: Profits do not enter the incentive compatibility constraints because the upfront transfer is not conditional on the behavior of individuals under the mechanism. In the supplementary material for this paper, we discuss an alternative version of the model in which expected tax revenues are not redistributed ex ante (before the operation of the mechanism), but ex post; that is, after θ has been observed, each individual receives a transfer

$$\frac{1-\rho}{n}\left(\sum_{i=1}^{n} t_i(\theta) - k(q(\theta))\right) .$$

In the supplementary material for this paper, it is shown that the above sequential structure can be imposed without loss of generality: The outcome of the model with a redistribution of expected tax revenues after the operation of the mechanism can be replicated by the model with an upfront transfer prior to the operation of the mechanism. The reason is that an ex post assignment of tax revenues is formally equivalent to an alternative specification of the payment rule $(t_i)_{i \in I}$ in a model in which expected tax revenues are distributed ex ante. Hence, the set of admissible allocations does not depend on this modeling choice.

An alternative specification of the participation constraints The individuals' share of tax revenue does not enter the participation constraints. This may be questioned on the following grounds. The participation constraints serve to ensure that individuals are not worse off as compared to a status quo situation with no public-goods provision. If their transfer income provides them with utility that they would not be able to realize in the status quo, then this should be included in the utility that they derive from public goods provision. Accordingly, the appropriate version of, say, the ex interim participation constraints would be as follows; for all i, and all θ_i,

$$\theta_i Q_i(\theta_i) - T_i(\theta_i) + \frac{1-\rho}{n} R^* \geq 0. \tag{7}$$

Again, it turns out that the specification with participation constraints as in (7) can be interpreted as a special case of the model with the participation constraints that do not include transfer payments. This is also clarified in the supplementary material for this paper.

5.2 The main result

We can now state the paper's main result in a formal way. The following Proposition compares, for an arbitrary individual i, ex ante expected utility with ex interim participation constraints, V_i^{int}, and ex ante participation constraints, V_i^{ant}. If V_i^{ant} is larger than V_i^{int} then the use of coercion is legitimate in the sense that the individual in question is, in an ex ante sense, made better off if coercion is possible.

Proposition 4 *Public goods provision is surplus-maximizing with ex ante participation constraints and distorted downwards with ex interim participation constraints. Moreover, there exists $\hat{\rho} \in (0,1)$ such that $V_i^{ant}(\rho) > V_i^{int}(\rho)$ if and only if $\rho < \hat{\rho}$.*

The Proposition shows that the legitimacy of coercion depends on the mechanism designer's degree of benevolence. If it is high, then expected payoffs with coercion are close to the symmetric Pareto-efficient mechanism. In this case, the imposition of ex interim participation constraints is harmful for individuals because it results in a lower level of aggregate surplus.

By contrast, for a low degree of benevolence, individuals prefer the imposition of ex interim participation constraints, even though this implies that public goods provision is inefficient. Given that the mechanism designer retains almost the whole surplus, they cannot benefit from surplus-maximizing public goods provision. The only remaining source of payoffs is therefore the information rent that individuals can reap provided that ex interim participation constraints are imposed. Hence, they prefer a larger fraction of a smaller, second-best surplus over a smaller fraction of the maximal, first best surplus.

This result can be summarized as follows. If public goods are provided in a benevolent way, the use of coercion is legitimate. A benevolent mechanism designer acts in the interests of individuals and should hence face as few constraints as possible. By contrast, if a malevolent institution is in charge of public goods provision then the use of coercion is not legitimate. A malevolent mechanism designer maximizes its own well-being at the expense of individuals. Hence, it is in the interest of individuals that he faces as many constraints as possible.

5.3 The main argument

To provide an intuitive understanding of the paper's main result in Proposition 4, we discuss the main steps in the proof in a semi-formal, heuristic way.

Ex interim participation constraints. Consider first the mechanism designer's problem with ex interim participation constraints. Once upfront payments to individuals are made, the mechanism designer aims at revenue maximization. Whatever his provision rule is, he will therefore choose the payments of individuals so that expected revenues $E\left[\sum_{i=1}^{n} t_i(\theta)\right]$ are maximized.

In the presence of incentive compatibility and ex interim participation constraints the maximal revenue equals the sum of the virtual valuations

$$E\left[\sum_{i=1}^{n} t_i(\theta)\right] = E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q(\theta)\right] .$$

The mechanism designer therefore chooses the provision rule which maximizes the second-best surplus defined as

$$E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q(\theta) - k(q(\theta))\right] .$$

We denote this provision rule in the following by q_R^{**}. With this second best provision rule, public goods provision is, for every θ, distorted downwards relative to the surplus-maximizing level. This follows since q_R^{**} is characterized implicitly by the first order condition,

$$\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) = k'(q_R^{**}(\theta)) ,$$

whereas the surplus-maximizing provision level is given by $\sum_{i=1}^{n} \theta_i = k'(q^*(\theta))$.

The expected payoffs of individuals, from the ex ante perspective are derived as follows: First, with ex interim participation constraints, they get an information rent which equals $E\left[\frac{1-P(\theta_i)}{p(\theta_i)} q_R^{**}(\theta)\right]$; second, they get a fraction $\frac{1-rho}{n}$ of the second-best profit. After algebraic manipulations which make use of the assumption that the individuals' taste parameters are *iid* random variables, we can therefore derive the following expression for individual i's expected payoff,

$$V_i^{int}(\rho) = \frac{1-\rho}{n} E\left[\sum_{i=1}^{n} \theta_i q_R^{**}(\theta) \ - \ k(q_R^{**}(\theta))\right] + \rho E\left[\frac{1-P(\theta_i)}{p(\theta_i)} q_R^{**}(\theta)\right] . \qquad (8)$$

Hence, an individual's expected payoff is a convex combination of the expected transfer and the information rent. In particular, the larger ρ, i.e., the smaller the fraction of revenue that is transferred to individuals, the larger is the contribution of the information rent to the individuals' expected payoff.

Ex ante participation constraints. The mechanism designer now aims at tax revenue maximization subject to ex ante participation constraints and incentive compatibility constraints.

However, to characterize the outcome of this mechanism design problem, we may ignore the incentive constraints. One can show that, if there is some mechanism that satisfies the budget constraint and guarantees individuals some non-negative expected utility level, then there is also another mechanism that generates these expected utility levels and is, in addition, incentive compatible.

Obviously, at a solution to the mechanism design problem, all ex ante participation constraints are binding. This implies that, for each individual i, $E[t_i(\theta)] = E[\theta_i q(\theta)]$. Upon substituting these expected payments in the mechanism designer's objective function, we find that he chooses q in order to maximize

$$E\left[\sum_{i=1}^{n} t_i(\theta) - k(q(\theta))\right] = E\left[\sum_{i=1}^{n} \theta_i q(\theta) - k(q(\theta))\right] \ ,$$

i.e., he chooses to provide public goods in a surplus-maximizing way.

Given that, with ex ante participation constraints, individuals are unable to reap an information rent, their expected payoff from the ex ante perspective consists entirely of their share of the first surplus which is given by

$$V_i^{ant}(\rho) = \frac{1-\rho}{n} E\left[\sum_{i=1}^{n} \theta_i q^*(\theta) \ - \ k(q^*(\theta))\right] \ . \tag{9}$$

Note that, whatever ρ, the mechanism designer chooses to provide public goods according to q^*, i.e., public goods provision is surplus-maximizing. Hence, the parameter ρ only affects the distribution of the surplus between the mechanism designer and the individuals, but has no bearing on the public goods provision rule. In particular, if ρ is close to 1, the expected payoff of individuals is close to zero.

Comparison. These derivations imply that

$$\begin{aligned} & V_i^{ant}(\rho) - V_i^{int}(\rho) \\ &= \tfrac{1-\rho}{n}\Big\{E\left[\textstyle\sum_{i=1}^{n} \theta_i q^*(\theta) \ - \ k(q^*(\theta))\right] - E\left[\textstyle\sum_{i=1}^{n} \theta_i q_R^{**}(\theta) \ - \ k(q_R^{**}(\theta))\right]\Big\} \\ & -\rho E\left[\tfrac{1-P(\theta_i)}{p(\theta_i)} q_R^{**}(\theta)\right] \ . \end{aligned}$$

The expression in curly brackets is the difference between the first-best and the second-best surplus, which is strictly positive. Hence, $V_i^{ant}(\rho) - V_i^{int}(\rho)$ is monotonically decreasing in ρ. Moreover it is easily verified that

$$V_i^{ant}(0) - V_i^{int}(0) > 0$$

and that

$$V_i^{ant}(1) - V_i^{int}(1) < 0 \ .$$

Consequently, there must exist an intermediate value $\hat{\rho}$ so that $V_i^{ant}(\rho) - V_i^{int}(\rho) \geq 0$ if and only if $\rho \leq \hat{\rho}$.

6 Concluding Remarks

The analysis has provided an answer to the question whether the financing of public goods should be subject to participation constraints. The advantage of a system based on participation constraints is that all individuals benefit from public goods provision. The disadvantage, however, is that public goods finance is generally insufficient to induce efficient outcomes. Which of these two forces is dominating depends on whether or not there are pronounced agency problems between individuals and the institution in charge of public goods provision. If the latter acts in the individuals' best interests, then the imposition on participation constraint is not attractive. By contrast, if it seeks to maximize his own payoff at the expense of individuals, then participation constraints should be imposed.

These results challenge the view that strong participation constraints are a "natural" ingredient of a model of public goods provision. The question whether they should be imposed depends on whether there is a tradeoff between distributive considerations (who gets how much of the surplus) and efficiency considerations (making the surplus as large as possible). This observation can serve as a starting point for future research. In particular, political economy models or models of public production and regulation that contain a more detailed description of institutions, could be used to study whether the resulting distribution of the surplus justifies the use of coercion for public goods finance.

References

Acemoglu, D., Golosov, M., and Tsyvinski, A. (2008). Markets versus governments. *Journal of Monetary Economics*, 55:159–189.

Adam, A., Delis, M., and Kammas, P. (2011). Are democratic governments more efficient. *European Journal of Political Economy*, 27:75–86.

Arrow, K. (1979). The property rights doctrine and demand revelation under incomplete information. In Boskin, M., editor, *Economics and Human Welfare*. Academic Press, New York.

Barro, R. (1973). The control of politicians: An economic model. *Public Choice*, 14:19–42.

Becker, S., Egger, P., and Seidel, T. (2009). Common policial culture: Evidence on regional corruption contagion. *European Journal of Political Economy*, 25:300–310.

Bergstrom, T., Blume, L., and Varian, H. (1986). On the private provision of public goods. *Journal of Public Economics*, 29:25–49.

Brennan, G. and Buchanan, J. (1980). *The Powet to Tax. Analytical Foundations of a Fiscal Constitution*. Cambridge University Press, New York, Michigan.

Buchanan, J. and Tullock, G. (1962). *The Calculus of Consent*. University of Michigan Press, Ann Arbor.

Coase, R. (1960). The problem of social cost. *Journal of Law and Economics*, 3:1–44.

d'Aspremont, C., Crémer, J., and Gérard-Varet, L. (2004). Balanced Bayesian mechanisms. *Journal of Economic Theory*, 115:385–396.

d'Aspremont, C. and Gérard-Varet, L. (1979). Incentives and incomplete information. *Journal of Public Economics*, 11:25–45.

Efendic, A., Pugh, G., and Adnett, N. (2011). Institutions and economic performance: A meta-regression analysis. *European Journal of Political Economy*, 27:586–599.

Fehrejohn, J. (1986). Incumbent performance and electoral control. *Public Choice*, 50:5–25.

Grüner, H. and Koriyama, Y. (2011). Public goods, participation constraints, and democracy: A possibility theorem. Mimeo, University of Mannheim.

Güth, W. and Hellwig, M. (1986). The private supply of a public good. *Journal of Economics*, Supplement 5:121–159.

Hellwig, M. (2003). Public-good provision with many participants. *Review of Economic Studies*, 70:589–614.

Hellwig, M. (2005). A utilitarian approach to the provision and pricing of excludable public goods. *Journal of Public Economics*, 89:1981–2003.

Hillman, A. and Ursprung, H. (2000). Political culture and economic decline. *European Journal of Political Economy*, 16:189213.

Holmstrom, B. and Myerson, R. (1983). Efficient and durable decision rules with incomplete information. *Econometrica*, 51:1799–1819.

Ledyard, J. and Palfrey, T. (1999). A characterization of interim efficiency with public goods. *Econometrica*, 67:435–448.

Lindahl, E. (1919). *Die Gerechtigkeit der Besteuerung.* Lund.

Mailath, G. and Postlewaite, A. (1990). Asymmetric bargaining procedures with many agents. *Review of Economic Studies*, 57:351–367.

Mas-Colell, A., Whinston, M., and Green, J. (1995). *Microeconomic Theory.* Oxford University Press, New York.

Myerson, R. and Satterthwaite, M. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 28:265–281.

Neeman, Z. (2004). The relevance of private information in mechanism design. *Journal of Economic Theory*, 117:55–77.

Persson, T. and Tabellini, G. (2000). *Political Economics: Explaining Economic Policy.* Cambridge, MA, MIT Press.

Rob, J. (1989). Pollution claim settlements under private information. *Journal of Economic Theory*, 47:307–333.

Samuelson, P. (1954). The pure theory of public expenditure. *Review of Economics and Statistics*, 36:387–389.

Wrede, M. (2001). Yardstick competition to tame the Leviathan. *European Journal of Political Economy*, 17:705721.

A Appendix

The proofs in the Appendix refer repeatedly to the characterization of incentive-compatible and tax-revenue-maximizing mechanisms in the supplementary for this paper. The supplementary material provides these characterizations for a model with discrete types. The continuous type space analogs of these characterizations are well-known, see for instance, Güth and Hellwig (1986), Hellwig (2005), or the textbook treatment Mas-Colell et al. (1995). For a reader who is familiar with these techniques the Appendix should therefore be self-contained. A reader who has not yet been exposed to them, may wish to consult the supplementary material first.

A.1 Proof of Proposition 1

Only if - part. We show that for every Pareto-efficient mechanism the budget constraint is binding and the provision rule is surplus-maximizing.

Without loss of generality of we can characterize a Pareto-efficient mechanism as the solution of the following optimization problem: Choose a mechanism in order to maximize $E[\theta_1 q(\theta) - t_1(\theta)]$ subject to the incentive compatibility constraints in (2), the budget constraint in (1) and the following set of reservation utility constraints: For each $i \neq 1$,

$$E[\theta_i q(\theta) - t_i(\theta)] \geq \bar{u}_i \ , \tag{14}$$

for some given vector of reservation utility levels $(\bar{u}_2, \ldots, \bar{u}_n)$.

Consider a relaxed problem which does not include the incentive compatibility constraints. It can be easily shown that the solution to this relaxed problem is such that the budget constraint and the constraints in (14) are binding. Moreover, public goods provision has to be surplus-maximizing. In the following we show that this solution can also be achieved subject to incentive compatibility constraints.

The surplus-maximizing provision rule satisfies $q(\theta_{-i}, \theta^l) < q(\theta_{-i}, \theta^{l+1})$, for all i and l. This implies that for all i, and all l, the monotonicity constrained $Q_i(\theta^{l+1}) > Q_i(\theta^l)$, is satisfied.

Given this provision rule, if we choose the expected payments of each individual i in such a way that all local downward incentive compatibility constraints are binding,

$$\theta^l Q_i(\theta^l) - T_i(\theta^l) = \theta^l Q_i(\theta^{l-1}) - T_i(\theta^{l-1}),$$

then Lemma 2 in the supplementary material implies that the resulting allocation is incentive compatible, and Lemma 3 implies that the expected utility of individual i is given by

$$E[\theta_i q(\theta) - t_i(\theta)] = E\left[\left(\frac{1 - P(\theta_i)}{p(\theta_i)}\right) q(\theta)\right] - T_i(\theta^0) \ .$$

Consequently, if we choose for each $i \neq 1$, $T_i(\theta^0)$ such that

$$E\left[\left(\frac{1 - P(\theta_i)}{p(\theta_i)}\right) q(\theta)\right] - T_i(\theta^0) = \bar{u}_i$$

and for all $l > 0$,

$$T_i(\theta^l) = \theta^l(Q_i(\theta^l) - Q_i(\theta^{l-1})) + T_i(\theta^{l-1}) \ ,$$

then the resulting allocation is incentive compatible and yields for each individual $i \neq 1$, the same expected utility as the solution of the relaxed problem. We proceed in a similar way with individual 1, except that the utility level for individual 1, u_1^*, follows from the solution of the relaxed problem and is given by

$$u_1^* = E\left[\sum_{i=1}^{n} \theta_i q^*(\theta) - k(q^*(\theta))\right] - \sum_{i=2}^{n} \bar{u}_i \ . \tag{15}$$

It remains to be shown that the payments are such that the budget constraint holds as an equality. By construction, $E[t_i(\theta)] = E[\theta_i q^*(\theta)] - \bar{u}_i$, for $i \neq 1$, and $E[t_1(\theta)] = E[\theta_1 q^*(\theta)] - u_1^*$. This implies that

$$E\left[\sum_{i=1}^{n} t_i(\theta)\right] = E\left[\sum_{i=1}^{n} \theta_i q^*(\theta)\right] - \left(u_1^* + \sum_{i=2}^{n} \bar{u}_i\right) \overset{(15)}{=} E\left[k(q^*(\theta))\right] \ .$$

If-Part. We now show that every incentive compatible mechanism such that the provision rule is surplus-maximizing and the budget constraint holds as an equality is Pareto-efficient.

The proof proceeds by contradiction. Suppose that $(q, t_1, \ldots, t_n)$ is an incentive compatible mechanism such that q is surplus-maximizing and the budget constraint holds as an equality. Suppose there exists an incentive compatible mechanism $(q', t'_1, \ldots, t'_n)$ such that, for all i,

$$E[\theta_i q'(\theta) - t'_i(\theta)] \geq E[\theta_i q(\theta) - t_i(\theta)] \ ,$$

with a strict inequality for some i. This implies that

$$E\left[\sum_{i=1}^{n}(\theta_i q'(\theta) - t_i'(\theta))\right] > E\left[\sum_{i=1}^{n}(\theta_i q(\theta) - t_i(\theta))\right] ,$$

Using that $E\left[\sum_{i=1}^{n} t_i(\theta)\right] = E[k(q(\theta))]$, and that $E\left[\sum_{i=1}^{n} t_i'(\theta)\right] \geq E[k(q'(\theta))]$, this implies that

$$E\left[\sum_{i=1}^{n}\theta_i q'(\theta) - k(q'(\theta))\right] > E\left[\sum_{i=1}^{n}\theta_i q(\theta) - k(q(\theta))\right] .$$

This contradicts the assumption that q is surplus-maximizing. ■

A.2 Proof of Proposition 2

Only if - part. By Proposition 1, under every Pareto-efficient mechanism the provision rule is equal to q^*. Given this provision rule, Lemma 7 in the supplementary material implies that the maximal revenue that is possible in the presence of ex interim participation constraints equals

$$E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q^*(\theta)\right] .$$

If this is smaller than $E[k(q^*(\theta))]$, budget balance can not be achieved. Hence, constrained efficiency can not be achieved.

If - part. We need to show that if

$$E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q^*(\theta)\right] \geq E[k(q^*(\theta))] .$$

then, given the surplus-maximizing provision rule q^*, $(t_1, \ldots, t_n)$ can be chosen such that the budget constraint binds and that for all i, the incentive compatibility constraints and the ex interim participation constraints are satisfied.

By the arguments in the proof of Lemma 7, the participation constraints of individual i are satisfied if and only if $T_i(\theta^0) \leq 0$. Since the provision rule q^* implies that the monotonicity constraints $Q_i(\theta^l) \geq Q_i(\theta^{l-1})$ are satisfied for all i and l, Lemmas 2 and 3 the supplementary material imply that incentive compatibility holds if expected payments are chosen such that all local downward incentive compatibility constraints are binding and that the expected payments of individual i are in this case equal to

$$E[t_i(\theta)] = E\left[\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q(\theta)\right] + T_i(\theta^0) .$$

Now choose

$$T_i(\theta^0) = \frac{1}{n}\left(E[k(q^*(\theta))] - E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right)q^*(\theta)\right]\right) ,$$

for all i. By assumption this is smaller or equal to zero, so that the ex interim participation constraints are satisfied, for all i. It remains to be shown that budget balance holds. This follows since, by construction,

$$\begin{aligned} E\left[\sum_{i=1}^{n} t_i(\theta)\right] &= E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right)q^*(\theta)\right] \\ &\quad + \left(E[k(q^*(\theta))] - E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right)q^*(\theta)\right]\right) . \end{aligned}$$

■

A.3 Proof of Proposition 3

Step 1. At a solution to the second best problem, the budget constraint has to be binding. Otherwise it would be possible to reduce the expected payments of individuals without violating any of the incentive compatibility or participation constraints, and without violating the budget constraint. Hence, at a solution to the second best problem

$$E\left[\sum_{i=1}(\theta_i q(\theta) - t_i(\theta))\right] = E\left[\left(\sum_{i=1}\theta_i\right)q(\theta) - k(q(\theta))\right] .$$

Step 2. The expected revenue $E[\sum_{i=1}^{n} t_i(\theta)]$ at a solution of the second best problem satisfies

$$E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right)q(\theta)\right] \geq E[\sum_{i=1}^{n} t_i(\theta)] .$$

This follows from the arguments in the proof of Lemma 7 in the supplementary material, which imply, that if, for a given provision rule q, $E[\sum_{i=1}^{n} t_i(\theta)]$ is maximized taking only a subset of the constraints of the second best problem – namely the ex interim participation constraints and the local downward incentive compatibility constraints – into account, then the maximal revenue equals $E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right)q(\theta)\right]$. This expression is therefore an upper bound on the expected payments of individuals. Combining this observation and the budget constraint yields

$$E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right)q(\theta)\right] \geq E[k(q(\theta))] . \tag{16}$$

Step 3. Steps 1 and 2 imply that the surplus that is generated at a solution of the auxiliary problem to maximize $E\left[\left(\sum_{i=1}\theta_i\right)q(\theta) - k(q(\theta))\right]$ subject to the constraint in (16) is an upper bound on the second best surplus. Moreover, it is straightforward to verify that (16) is binding if and only if

$$E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right)q^*(\theta)\right] < E[k(q^*(\theta))] \ .$$

Step 4. Suppose that (16) is not binding. Then provision rule q^* and the payments in the mechanism in the proof of the if-part of Proposition 2 solve the auxiliary problem. Moreover, with this mechanism the surplus of public goods provision is shared equally among individuals, i.e., for all i,

$$E[\theta_i q(\theta) - t_i(\theta)] = \frac{1}{n}\,E\left[\sum_{i=1}^{n}\theta_i q^*(\theta) - k(q^*(\theta))\right] \ .$$

This proves the first statement in Proposition 3.

Step 5. Now suppose that (16) is binding. The provision rule that solves the auxiliary problem satisfies the monotonicity constraint, $Q_i(\theta^l) \geq Q_i(\theta^{l-1})$, for all i, and l. This follows because the optimal level of $q(\theta)$ is either equal to zero or given by the first order condition,

$$k'(q(\theta)) = \sum_{i=1}^{n}\theta_i - \frac{\lambda}{1+\lambda}\sum_{i=1}^{n}\frac{1-P(\theta^i)}{p(\theta_i)}$$

where λ is the multiplier on the constraint. The assumption that the hazard rate is decreasing implies that whenever, for one individual the taste parameter θ^l is replaced by the taste parameter θ^{l+1}, the right hand side goes up, which implies that $q(\theta_{-i}, \theta^l) \leq q(\theta_{-i}, \theta^{l+1})$, for all i, θ_{-i}, and l. This implies, in particular, that $Q_i(\theta^l) \leq Q_i(\theta^{l-1})$, for all i, and l. This follows from Lemma 7 in the supplementary material. This implies that the the surplus that is generated by the auxiliary problem can be achieved by a second best mechanism. Moreover, the arguments in the proof of this Lemma show that this requires that for all i, the ex interim Participation constraint $T_i(\theta^0) \leq 0$ and all local downward incentive compatibility constraints are binding. Lemma 3 in the supplementary material implies that, at a solution to the second best problem, for all i, ex ante expected utility is equal to

$$E[\theta_i q(\theta) - t_i(\theta)] = E\left[\frac{1-P(\theta_i)}{p(\theta_i)}q^{**}(\theta)\right] \tag{17}$$

where $q^{**}(\theta)$ is the provision rule that solves the auxiliary problem. Given that the constraint of the auxiliary problem is binding we have

$$E\left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q^{**}(\theta)\right] = E[k(q^{**}(\theta))] \ ,$$

or, equivalently,

$$E\left[\sum_{i=1}^{n}\frac{1-P(\theta_i)}{p(\theta_i)} q^{**}(\theta)\right] = E\left[\left(\sum_{i=1}^{n}\theta_i\right) q^{**}(\theta) - k(q^{**}(\theta))\right]$$

Using that the random variables $(\theta_i)_{i=1}^n$ are independently and identically distributed, this implies that

$$E\left[\frac{1-P(\theta_i)}{p(\theta_i)} q^{**}(\theta)\right] = \frac{1}{n}\, E\left[\left(\sum_{i=1}^{n}\theta_i\right) q^{**}(\theta) - k(q^{**}(\theta))\right] \tag{18}$$

Equations (17) and (18) imply that at a solution to the second best problem,

$$E[\theta_i q(\theta) - t_i(\theta)] \quad = \quad \frac{1}{n}\, E\left[\left(\sum_{i=1}^{n}\theta_i\right) q^{**}(\theta) - k(q^{**}(\theta))\right] \ .$$

Since $q^{**} \neq q^{*}$, this is less than the ex ante expected payoff under the symmetric constraint efficient mechanism. ■

A.4 Proof of Proposition 4

The proof is based on Lemmas 1 - 7 in the supplementary material, which provide a characterization of incentive compatible mechanisms and of revenue-maximizing mechanisms.

The optimal mechanism with ex interim participation constraints. It follows from Lemma 5, that, at a solution to the mechanism design problem, for all i, the participation constraints in (3) are binding for $\theta_i = \theta^0$ and is slack otherwise. Otherwise it would be possible to increase the policy-maker's revenue while holding the policy-maker's provision rule fixed.

Consider the relaxed problem of choosing $(q, t_1, \ldots, t_n)$ in order to maximize R subject to the downward incentive compatibility constraints, for any $l > 0$,

$$\theta^l Q_i(\theta^l) - T_i(\theta^l) \geq \theta^l Q_i(\theta^{l-1}) - T_i(\theta^{l-1}) \ . \tag{19}$$

and the ex interim participation constraints in (3). It follows from Lemma 6 that at a solution to this problem, all downward incentive compatibility constraints are binding, and the participation constraint in (3) is binding for $\theta_i = \theta^0$ and are slack otherwise. Otherwise it would be possible to increase the policy-maker's revenue while holding the policy-maker's provision rule fixed. But then the arguments in the proof of Lemma 7 imply that the expected revenue of the policy-maker, at a solution to the relaxed problem, is equal to $E\left[\sum_{i=1}^n t_i(\theta)\right] = E\left[\sum_{i=1}^n \left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q(\theta)\right]$. Hence, the provision rule which is part of the solution of the relaxed problem maximizes

$$R = E\left[\sum_{i=1}^n \left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q(\theta) - k(q(\theta))\right] .$$

In the following this provision rule is denoted by q_R^{**}.

The relaxed problem takes only a subset of all incentive compatibility constraints into account. Hence, the expected profits that are generated by the mechanism which solves the relaxed problem are an upper bound on the expected profits that are generated by the mechanism that solves the "full" problem of maximizing R subject to all participation and incentive compatibility constraints.

It follows from Lemma 2 that if the provision rule q_R^{**} is such that the monotonicity constraints $Q_i(\theta^l) \geq Q_i(\theta^{l-1})$ are satisfied for all i and l, then the solution to the relaxed problem satisfies all incentive compatibility constraints and is hence also a solution to the full problem.

In the remainder we verify that under q_R^{**} the monotonicity constraints are indeed satisfied. For every given θ, $q_R^{**}(\theta)$ is either given by the first order condition,

$$\sum_{i=1}^n \left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) = k'(q_R^{**}(\theta)) , \tag{20}$$

or, if this equation has only a negative solution, equal to 0. Given that $h^l < h^{l-1}$, for all l, for each i, the left-hand side of the first order condition is strictly increasing in θ_i. Given that k is increasing and convex, this implies that, for each i and each l, $q_R^{**}(\theta_{-i}, \theta^l) < q_R^{**}(\theta_{-i}, \theta^{l+1})$, and, as a consequence, the monotonicity constraint $Q_i(\theta^{l+1}) > Q_i(\theta^l)$ holds for all i, and all l.

Given that all local downward incentive compatibility constraints and the participation constraints for $\theta_i = \theta^0$ are binding for all i (so that $T_i(\theta^0) = 0$), it follows from Lemma 3 that $E[\theta_i q(\theta) - t_i(\theta)] + \frac{1-\rho}{n} R^*$ is equal to

$$E\left[\frac{1-P(\theta_i)}{p(\theta_i)} q^{**}(\theta)\right] + \frac{1-\rho}{n} E\left[\sum_{i=1}^n \left(\theta_i - \frac{1-P(\theta_i)}{p(\theta_i)}\right) q^{**}(\theta) - k(q^{**}(\theta))\right]$$

Exploiting that the random variables $(\theta_i)_{i=1}^n$ are *iid* and rearranging term yields the expression for expected payoffs of individuals in equation (8).

The optimal mechanism with ex ante participation constraints We first consider a relaxed problem of maximizing R taking only the ex ante participation constraints in (4) into account. Obviously, this implies that, for each i, the participation constraint has to be binding, for each i, $E[t_i(\theta)] = E[\theta_i q(\theta)]$. Using this expression to substitute for $E[t_i(\theta)]$ in the definition of R yields $R = E\left[\left(\sum_{i=1}^n \theta_i\right) q(\theta) - k(q(\theta))\right]$. Hence, q^* is the tax-revenue-maximizing provision rule.

In the following we show that the there is a mechanism which is payoff equivalent to the solution of this relaxed problem and satisfies all incentive compatibility constraints.

The surplus-maximizing provision rule q^* is such that for each i, $Q_i(\theta^l) \leq Q_i(\theta^{l+1})$. If all local downward incentive compatibility constraints hold as an equality, then all incentive compatibility constraints are satisfied. This follows from Lemma 2. To complete the proof it suffices show that, given public goods provision according to q^*, there is a payoff equivalent mechanism which is such that all local downward incentive compatibility constraints hold as an equality.

If all local downward incentive constraints are binding, then the arguments in the proof of Lemma 2 imply that $T_i(\theta^l) = \theta^l Q_i(\theta^l) - \sum_{k=0}^{l-1} Q_i(\theta^k) + T_i(\theta^0)$, for all l, and hence

$$E[t_i(\theta)] = E\left[\left(\theta_i - \frac{1 - P(\theta_i)}{p(\theta_i)}\right) q^*(\theta)\right] + T_i(\theta^0)\,. \tag{21}$$

Now if we let, in addition, $T_i(\theta^0) = E\left[\frac{1-P(\theta_i)}{p(\theta_i)} q^*(\theta)\right]$, then also the ex ante participation constraints of all individuals are binding. This also implies that for each individual $E[\theta_i q^*(\theta) - t_i(\theta)] = 0$ and expected profits are equal to $R^* = E\left[\left(\sum_{i=1}^n \theta_i\right) q^*(\theta) - k(q^*(\theta))\right]$. The ex ante expected payoff of individuals consists entirely of the fraction of profits that they receive. This observation yields the expression for expected utility in equation (9).

Comparison of expected utilities. Using equations (8) and (9) we find that $V_i^{ant}(1-\rho) - V_i^{int}(1-\rho)$ is equal to

$$(1-\rho)\frac{1}{n}\left(E\left[\left(\sum_{i=1}^n \theta_i\right) q^*(\theta) - k(q^*(\theta))\right] - E\left[\left(\sum_{i=1}^n \theta_i\right) q_R^{**}(\theta) - k(q_R^{**}(\theta))\right]\right)$$
$$-\rho E\left[\left(\frac{1 - P(\theta_i)}{p(\theta_i)}\right) q_R^{**}(\theta)\right]$$

Since q^* maximizes the surplus from public goods provision, this expression is decreasing in ρ. Moreover, it is negative for ρ close to 1 and positive for ρ close to 0. ∎