Essays on Public Goods Provision and Income Taxation

Inauguraldissertation

zur Erlangung des akademischen Grades

eines Doktors der Wirtschaftswissenschaften

der Universität Mannheim

vorgelegt von

Diplom-Volkswirt Felix Bierbrauer

Mannheim 2006

Referent: Prof. Martin Hellwig, Ph.D.

Korreferent: Prof. Dr. Eckhard Janeba

Dekan: Prof. Konrad Stahl, Ph.D.

Tag der mündlichen Prüfung: 22. Mai 2006 *to Carina*

Acknowledgements

I am deeply grateful for the support of my supervisor, Martin Hellwig. In numerous discussions he questioned preliminary versions of this work and encouraged further efforts. I learned very much from his comments, his suggestions and most notably from his insistence on a conceptually sound analysis. Not least importantly, during a more cumbersome phase it was of great help that he remained optimistic with respect to the prospects of this dissertation.

I am also indebted to Eckhard Janeba for his encouragement of this work and for being my second thesis advisor. I thank the professors at the CDSEM at Mannheim University for an excellent graduate education that allowed me to acquire the skills that were a prerequisite for the completion of this work. My fellows at the CDSEM and in the Department of Economics provided me with a lively and inspiring atmosphere.

I also thank my colleagues at the Max Planck Institute for Research on Collective Goods in Bonn. The institute provides an excellent research infrastructure. Moreover, I enjoyed many interesting discussions with the researchers at the institute.

Not least of all, I benefited a lot from the collaboration with my friend and coauthor Marco Sahm. The interaction with Marco was both productive and a great pleasure.

I am very grateful to my parents, Peter and Ursula Bierbrauer. The reasons include their continuous support during my graduate education. My wife, Carina, is the person whose well-being was most heavily affected by this dissertation. Nevertheless she always encouraged my scientific work with patience and confidence. Her caring empathy was essential for the completion of this work.

Felix Bierbrauer

Contents

Chapter 1

Preface

This dissertation belongs to the field of normative public economics. It is concerned with the characterization of schemes of taxation and public good provision that are optimal from a welfare perspective.

More precisely, optimal tax systems and optimal rules for public good provision are analyzed based on the assumption that individuals are privately informed about their valuation of the public good and that an optimal rule for public good provision reflects this information; e.g. a larger average valuation of the public good among individuals in the economy implies that the quantity of the public good that is provided under an optimal rule goes up. As a consequence, an optimal provision rule has to be based on some procedure of information aggregation that allows this information to be acquired. The tax system is a key determinant for the task of information aggregation. If individuals are asked to communicate their valuation of the public good to "the system", they compare the utility gain from a larger level of public good provision to the utility burden that results from the need to generate larger tax revenues to cover the costs of public good provision. Individuals will hence communicate their "true" valuation of the public good only if these two forces are commensurate.

To illustrate this, consider a tax system that exempts individuals with a very low level of income from any tax payment. For these individuals an increased tax revenue requirement does not cause any utility burden. Hence those individuals welcome any increase in the quantity of the public good and are thus inclined to exaggerate when asked about their valuation of the public good. Alternatively, suppose that tax revenues are used not only to cover the cost of public good provision but also to finance direct income transfers to poor individuals. In this case, individuals with a rather high level of income might claim an excessive taste for public good provision just in order to ensure that the fraction of tax revenues that is dedicated to the transfer system does not become too large.

These considerations demonstrate that the problem of finding an optimal tax system is intertwined with the problem of finding an optimal rule for an informed decision on public good provision. An analysis based on the presumption that information on public goods preferences just happens to be available is too naive. Individuals might refuse to reveal this information because a higher level of public spending affects their personal tax bill. This concern is the topic of this dissertation. Each chapter contains a characterization of optimal tax schemes and provision rules under the premise that information on public goods preferences needs to be acquired.

A theory of optimal taxation and public good provision in the presence of uncertainty about public goods preferences is of value for two reasons. First, it is desirable to have a more complete theory, and below I will argue that such a theory has not yet been developed. Second, and more substantively, the use of tax revenues to cover the cost of public good provision and the need to assess the desirability of a public project prior to the final spending decision are real-world phenomena. A normative theory thus provides a benchmark that allows judgement about how real-world institutions deal with this problem.

The theory identifies outcomes that take specific constraints into account and are optimal under a given welfare function. Institutional constraints determine the set of tax instruments used for public goods finance. Technological constraints enter the analysis via the cost of public good provision that determines the tax revenue requirement in the public sector budget constraint. Finally, there are informational constraints. Individuals are privately informed about their public goods preferences. Hence, an optimal policy can use only those pieces of information that individuals are indeed willing to reveal to the system. Considerations of political feasibility enter the analysis via this latter set of constraints. Information can be acquired only if it is used in a way that is in line with the interests of individuals. These interests in turn are shaped by the tax system and the provision rule for public goods. The derivation of a normative benchmark that takes all these constraints into account is of limited use when it comes to recommendations for actual public policy. However, it provides a better understanding of their interplay and of the restrictions that become effective even under an ideal tax system. For instance, chapter 4 of this dissertation identifies a tradeoff between the desire to have an optimal redistributive tax system for a *given* level of public good provision and the problem of acquiring the information that is needed to determine the *optimal* quantity of the public good. It is shown that these two tasks can not be achieved simultaneously; that is, even the best policy-maker is not able to escape this problem.

In more abstract terms, a theory of optimal taxation, public good provision and information acquisition yields a characterization of constrained efficient allocations. In addition, with a given welfare assessment, *optimal* constrained efficient allocations can be studied. As an example, it is shown in chapter 4 that a constrained efficient utilitarian tax system displays a complementarity between the extent of redistribution and the decision on public good provision – relative to a situation where informational constraints are not taken into account.

1.1 Related Literature

The topics of optimal taxation and public good provision, on the one hand, and the elicitation of preferences for public goods, on the other hand, have so far been treated in separate branches of economic theory. For brevity, I will henceforth refer to the first branch of the literature as the *theory of optimal taxation* and to the second branch as the *theory of mechanism design*.

The setup in the theory of optimal taxation is as follows. A benevolent planner chooses several policy variables on the basis of some welfare function. The policy variables include various direct or indirect tax instruments and provision levels of public goods. The planner takes into account that individual consumption choices and labour supply decisions respond to the chosen policy and that a public sector budget constraint has to be satisfied.

In this theory the optimal quantity of a public good is determined according to a modified *Samuelson rule*, after Samuelson (1954). The optimal quantity is such that the sum of marginal valuations of the public good is equal to the marginal cost of public funds under the given tax system.[1] Obviously,

[1]There are numerous contributions to this literature, which differ according to the tax instruments that are used for public goods finance. Examples include Atkinson and Stern (1974), Wilson (1991), Boadway and Keen (1993), Nava et al. (1996), Sandmo (1998),

an application of the modified Samuelson rule requires the assumption that the sum of marginal valuations of the public good is commonly known.

The theory of mechanism design, by contrast, studies the problem of how to acquire this information.[2] In this theory, uncertainty about the sum of marginal valuations arises in the following way: individuals have private information on their own valuation of the public good and, as a consequence, the sum of all individual valuations or, equivalently, the *average valuation* of the public good is an unknown variable. In order to acquire this information, a mechanism designer asks all individuals to report their valuation.

The main question for the mechanism design approach is to what extent the classical *free rider problem* in public good provision may be resolved, i.e. it is concerned with the welfare costs of having to finance public good provision in such a way that any individual is willing to reveal her valuation of the public good.[3] In this framework an optimal payment scheme has a Pigouvian spirit in the sense that it forces each individual to internalize the consequences of her own preference announcement on the well-being of all other individuals in the economy.

Mechanism design theory differs from the theory of optimal taxation not only because it incorporates a problem of information aggregation. Further differences are the following. Mechanism design problems are typically based on

Hellwig (2005b, 2004) and Gaube (2000, 2005).

[2]This line of research starts with the invention of Clarke-Groves mechanisms by Clarke (1971) and Groves (1973).

[3]The focus of the early literature is the question whether one may have simultaneously ex-post efficiency and incentive compatibility. It has been addressed with two different solution concepts. Green and Laffont (1977) establish the impossibility of implementation in dominant strategies, while Arrow (1979) and d'Aspremont and Gérard-Varet (1979) establish the possibility of implementing an efficient allocation as a Bayesian Nash equilibrium.

environments in which all individuals have quasi-linear preferences, i.e. the marginal disutility of having to pay for the public good is constant for all individuals. This implies that taxation is not distortionary, in the sense of driving a wedge between marginal rates of substitution and marginal rates of transformation. Moreover, the direct or indirect tax instruments on which the theory of optimal taxation is based do not enter the analysis. Finally, the recent literature in this field incorporates participation constraints.[4] The literature on optimal taxation typically does not include such a restriction and takes the state's authority to rely on coercion as given.

This brief overview of the existing literature can be summarized as follows: There exists a normative theory that studies optimal tax systems and optimal decisions on public good provision under the premise that these two policy variables are linked through a public sector budget constraint. It does not, however, include a problem of preference elicitation. This concern is treated in a different branch of the literature, which does not consider the use of direct and indirect tax instruments for the purpose of public goods finance.

At a conceptual level the main objective of this dissertation is to provide a link between these two approaches that allows the study of the problems of optimal taxation, public good provision and information aggregation simultaneously.

[4] It starts out from the observation that ex-post efficiency, Bayesian incentive compatibility and voluntary participation are incompatible; see Güth and Hellwig (1986) and Mailath and Postlewaite (1990). The properties of second best allocations under incentive as well as participation constraints are analyzed by Schmitz (1997), Hellwig (2003) and Norman (2004).

1.2 Conceptual Issues

The attempt to introduce a problem of information aggregation into the theory of optimal taxation faces a conceptual problem. It is due to the fact that this theory typically analyzes a large economy with a continuum of individuals.[5] To see this, suppose that a large number of individuals is involved in a process of information aggregation that is used to determine the average valuation of the public good in the economy. The fact that the economy is large implies that no single individual has a direct impact on the outcome of this procedure. Hence, no individual has a motive for hiding his true valuation of the public good.

This reasoning is based on the notion of individual incentive compatibility that is typically used in the theory of mechanism design. Incentives are only needed for individuals who have some chance of being pivotal for the decision on public good provision. If the probability of being pivotal vanishes as one moves to a large economy, so does the need to specify appropriate incentives. This yields the conclusion that, in a large economy, information aggregation is not an incentive problem.

The work presented in the subsequent chapters of this dissertation takes a different view. It is based on the observation that the tax system shapes individual interests with respect to the decision on public good provision and that individuals who act according to these interests may refuse to reveal

[5]See Dierker and Haller (1990) for a discussion. A main reason for the consideration of a large economy in the theory of optimal income taxation is the critique of Piketty (1993) and Hamilton and Slutsky (2005). Accordingly, in a finite economy with a known cross-section distribution of characteristics, first best utilitarian redistribution can be achieved if one does not rely on an income tax but uses a more sophisticated game form. In a continuum economy, however, this problem does not arise if an appropriate version of the *law of large numbers for large economies* is assumed to hold, Guesnerie (1995).

their private information.

To make this more precise, suppose that a tax setting planner collects individual statements on public goods preferences in order to learn the average valuation. For the sake of concreteness, consider an individual with a modest valuation of the public good but a rather high tax bill. Suppose that this individual hopes that the average valuation of the public good turns out to be very low and that, as a consequence, the public good is not installed. The individual in question is thus inclined to understate her valuation of the public good and to contribute thereby contribute to the perception that the average valuation is lower than it actually is.

This perspective yields the conclusion that appropriate incentives are needed to prevent a misrepresentation of public goods preferences. Below I discuss two different approaches with which it is possible to tackle this conceptual problem. They both take the specific interests of individuals under a tax system into account and treat information aggregation as an incentive problem, even in a large economy.

Chapter 2 draws on literature in the field of political economy that discusses voting mechanisms as an instrument for information aggregation in the tradition of the Condorcet Jury Theorem.[6] This literature is based on the idea that sincere voting behavior – in which each individual votes for his preferred alternative – implies that the distribution of votes contains information about the distribution of preferences in the economy. Consequently, the outcome of a voting decision reflects this information. Chapter 2 draws an analogy to this literature. Incentive constraints are introduced into the

[6] See, for instance, Feddersen and Pesendorfer (1997). Piketty (1999) provides a survey of related literature.

analysis which ensure that it is in each individual's interest to behave sincerely. This allows the study of the properties of an optimal rule for public good provision under the constraint that individuals are willing to engage in informative voting about the level of public good provision.

Chapter 2 has no axiomatic foundation but is based on an analogy to a branchof the literature in political economics. Chapter 3 follows a more rigorous approach and introduces the notion of a *collectively incentive-compatible tax system*. Chapters 4 and 5 contain applications of this solution concept to the problem of optimal income taxation.

Collective incentive considerations are based on the idea that individuals may form coalitions in order to manipulate the perception of the average valuation of the public good and thereby the decision on public good provision. This circumvents the problem that, in a large economy, no single individual has a reason to behave strategically. Via coalition formation individuals can affect the outcome of a revelation mechanism even in a large economy. Consequently, an implementable rule for the use both of tax instruments and the level of public good provision has to be such that no coalition of individuals has an incentive to engage in a collective misrepresentation of public goods preferences. A tax system with this property is said to be *collectively incentive compatible.*[7]

The remaining part of this introduction makes this more concrete and gives an overview of the specific models and the results that are derived in the subsequent chapters.

[7]This definition of *collective incentive compatibility* is inspired by the notion of a *coalition-proof Nash equilibrium* introduced by Bernheim et al. (1986).

1.3 Chapter 2: Informative Voting and the Samuelson Rule

Chapter 2 is based on a joint research project with Marco Sahm from the Ludwigs-Maximilian University in Munich. It studies an environment with the following properties. The economy is large and each individual has a quasi-linear utility function that depends on the level of public good provision and the individual's contribution to the cost of public good provision. In this model *equal cost sharing* is the only budgetary feasible and individually incentive-compatible scheme of public goods finance.[8]

An individual's willingness to pay for the public good depends on two parameters, a binary taste parameter which indicates either a high or a low taste and, in addition, an ability parameter. Ceteris paribus, an individual's willingness to pay is an increasing function of the taste parameter and a decreasing function of the ability level. The latter effect reflects the idea that less able individuals suffer from a larger utility loss if forced to generate the income that is needed in order to meet a given payment obligation.[9]

The problem of information aggregation results from the assumption that there is uncertainty about the average valuation of the public good. This uncertainty is induced via the share of individuals with a high taste parameter, which is taken to be a random quantity. Its actual realization has to be deduced from the collection of individual statements on taste parameters in

[8]This follows from the assumptions that individuals possess private information on their personal characteristics and that the final decision on public good provision must depend only on the empirical distribution of characteristics, not on the personal characteristics of specific individuals. As a consequence, the payment scheme cannot differentiate between individuals with different characteristics.

[9]This idea is familiar from the theory of optimal income taxation.

a revelation game.

If a rule for public good provision only has to meet the requirements of feasibility and individual incentive compatibility, then the *optimal* rule for public good provision is a version of the Samuelson rule that equates the average willingness to pay for the public good when costs are shared equally with the marginal cost of the public good.

However, under this version of the Samuelson rule information on individual characteristics may be used in a way that runs counter to the interests of those individuals. To see this, consider an individual with a high taste realization but a very low skill level and, for the sake of concreteness, suppose this individual expects the state of the economy to be such that his own willingness to pay is below the average. Consequently, under the Samuelson rule, this individual expects the level of the public good to be too large. If the individual chooses his taste announcement while taking his preferred outcome of the revelation game into account, she should announce a low taste realization and thereby *"contribute"* indirectly to a more preferred perception of the state of the economy.

The notion of *informative voting* is introduced into the analysis to avoid outcomes with this property. It is borrowed from the field of political economy that analyzes how voting mechanisms incorporate information on the distribution of preferences. To make use of these ideas, any taste announcement is interpreted as a vote, i.e. a high (low) taste announcement is regarded as a vote in favor of a high (low) level of public good provision. It is assumed that individuals vote sincerely; that is, an individuals votes in favor of a high level of public good provision only if he indeed benefits from such an outcome.[10] This assumption imposes a constraint on the task of information aggregation. Information on the average taste level becomes available only if the public good is provided in such a way that all individuals with a high (low) taste realization prefer a high (low) provision level over a low (high) provision level.

The main result of the analysis in Chapter 2 is a characterization of the optimal utilitarian provision rule which satisfies these constraints. It is shown that the optimal extent of information aggregation is inversely related to a specific measure of preference polarization in the economy.[11] To construct this measure, the economy is divided into two groups: those individuals with a high taste parameter and those individuals with a low taste parameter. Heterogeneity with respect to ability levels implies that there is *within-group polarization*: an individual with low skills and a high taste parameter has a willingness to pay for the public good that is small relative to the one of an individual with high skills and a high taste parameter.

The analysis shows that the larger the degree of *within-group polarization*, the smaller the sensitivity of an optimal provision rule to changes in the average willingness to pay for the public good. Put differently, more polarization implies that an optimal provision rule uses less information on public goods preferences.

[10] The term *sincere voting* is used in the field of political economy. Austen-Smith and Banks (1996) discuss this terminology more extensively.

[11] These observations are similar in spirit to classical results from the signalling literature. See e.g. Crawford and Sobel (1982), Schultz (1996), Grossman and Helpman (2001, Ch. 4).

1.4 Chapter 3: Collectively Incentive Compatible Tax Systems

Chapter 3 introduces a more general framework for the joint study of optimal tax systems and the revelation of preferences for public goods. Preferences are not assumed to be quasi-linear as in Chapter 2. However, the quasi-linear environment is a special case of the setup in Chapter 3.

More precisely, Chapter 3 introduces a problem of information aggregation into the model that is typically used in the theory of optimal income taxation: Individuals have private information on their earning ability. Simultaneously it is assumed that the empirical cross-section distribution of earning ability is commonly known; i.e. with respect to earning ability there is no uncertainty at the aggregate level. The new assumptions introduced in Chapter 3 are that, in addition, individuals have private information on their valuation of a public good. Moreover, this uncertainty about individual valuations does not wash out in the aggregate. There is aggregate uncertainty because the joint cross-section distribution of earning ability and valuations of the public good is not commonly known.

The characterization of admissible tax systems and provision rules for public goods is treated as a problem of mechanism design. For this purpose, tax systems are identified with the set of *decentralizable* allocations that has been defined by Hammond (1979). For such an allocation, there exists a tax system such that the commodity bundle of each individual is the solution of a standard consumer choice problem. For instance, if individuals care only about consumption and leisure and there is no issue of information aggregation, then the problem of finding an optimal decentralizable allocation is equivalent to the problem of finding in optimal income tax, in the sense of

Mirrlees (1971).

The notion of a *collective incentive-compatible tax system* addresses the incentive issues that come into play with the elicitation of public goods preferences. As has been explained above, these incentive constraints ensure that a planner who decides on the use of tax instruments and on public good provision does not rely on the availability of private information that individuals are not willing to reveal once they are given the opportunity to engage in manipulative collective action.

The main formal result in this chapter demonstrates that considerations of *individual* incentive compatibility and *collective* incentive compatibility can be analyzed separately if the preferences of individuals have a certain structure. If the utility contribution of the public good is additively separable from the utility contribution of private goods, the following can be established: *Collective incentive compatibility* holds if no coalition of individuals benefits from a manipulation of public goods preferences, taking as given that these individuals reveal their earning ability. The revelation of earning ability is ensured by *individual* incentive compatibility constraints. Consequently, there is no need to worry about coalitions that manipulate the profile of earning abilities.

While this is per se not a deep insight, it proves convenient for a more explicit characterization of implementable allocations in more specific environments. To illustrate this, the quasi-linear economy of Chapter 2 is once again addressed. In this environment, an application of the separability result shows that – under certain assumptions concerning the process of information aggregation – the set of allocations that are collectively incentive compatible coincides with the set of allocations that are admissible under the requirement of informative voting in Chapter 2. This is notable for two reasons.

First, it shows, for a specific environment, that the set of collectively incentive compatible allocations can be explicitly characterized. Second, it implies that the requirement of informative voting can be equivalently interpreted as a condition that precludes the formation of coalitions by individuals with the same taste realization.

1.5 Chapter 4: Optimal Income Taxation and Public Good Provision in a Two-Class Economy

In Chapter 4 the problem of preference elicitation is introduced into a model of optimal utilitarian income taxation. This literature is concerned with the *equity-efficiency tradeoff* that arises in the following way. Heterogeneity with respect to ability levels generates a utilitarian desire to redistribute consumption from "the rich" class of individuals to the "poor" class. First best utilitarian welfare, however, is out of reach because individual ability levels are private information. An optimal tax system distorts labour supply decisions in order to realize welfare gains from redistribution. These distortions are the source of the *equity-efficiency tradeoff*.

Chapter 4 studies a *two-class economy* in which individuals either have a high or a low level of earning ability.[12] Uncertainty about public goods preferences is introduced in the following way: Valuations of the public good are either high or low. Moreover, these public goods preferences are assumed to be perfectly correlated with earning ability; that is, all individuals with

[12] This environment has received some attention in the literature on optimal taxation; see e.g. Mirrlees (1975), Stiglitz (1982, 1987), Boadway and Keen (1993), Nava et al. (1996) or Gaube (2005).

the same level of earning ability also have the same valuation of the public good. With this specific information structure, the problem of *information aggregation* is concerned with the elicitation of the public goods preferences of individuals with high and low ability, respectively.[13]

The solution concept used to address these two problems is the *collectively incentive compatible tax-system* introduced in Chapter 3. In the two-class economy collective incentive compatibility holds if individuals with same level of ability – i.e. those individuals who belong to same class – do not benefit from a joint collective lie on their taste parameter. Introducing these additional constraints into the problem of optimal utilitarian income taxation allows the study of how the problem of preference elicitation interacts with the *equity-efficiency tradeoff*.

The interaction arises because of the fact that an optimal utilitarian income tax creates conflicting views of the desirability of public good provision. It is shown in Chapter 4 that the need to generate tax revenues for the public good affects more able and less able individuals differently. It is a consequence of the *equity-efficiency tradeoff* that the utility burden from the cost of public good provision is larger for the less productive individuals. Whenever resources are needed for the public good, incentive compatibility constraints prevent the planner from extracting larger tax payments from the more able class.[14]

Now suppose that individuals of either class are asked to reveal their true

[13] The information structure is such that taste realizations and ability levels are binary variables. This implies a similarity to the two-dimensional screening models of Armstrong and Rochet (1999) and Cremer et al. (2001).

[14] These properties are established by Weymark (1986, 1987) for an optimal income tax model with a finite number of different classes and preferences that are quasi-linear in leisure.

taste realization. The fact that the utility burden of the larger tax revenue requirement is felt more by the less able class creates the following pattern of collective incentive problems. More able individuals tend to claim an excessive desire for public good provision, and less able individuals are inclined to understate the desirability of provision.

A decision on provision that reflects the "true" aggregate valuation of the public good necessitates an adjustment of the transfer system that corrects these biases. It is shown in Chapter 4 that an *optimal* elimination of these collective incentive problems is characterized by a *complementarity* between the level of redistribution and the decision on public good provision, relative to an *equity-efficiency tradeoff*, without a problem of information aggregation. To prevent the more productive class from exaggerating, public good provision has to be accompanied by an increased level of redistribution. Similarly, the less productive are prevented from understating their valuation of the public good by a reduced level of redistribution if there is no public good provision.

1.6 Chapter 5: Distortionary Taxation and the Free-Rider Problem

Chapter 5 is based on the same environment as Chapter 4; that is, there is a two-class economy with uncertainty about the public goods preferences of more able and less able individuals, respectively.

The difference from Chapter 4 lies in the tax instrument that is used for the purpose of public goods finance. Chapter 5 assumes that a linear tax on income is raised solely to cover the cost of public good provision. This specification excludes the interaction between the transfer system and the problem of preference elicitation that arises under non-linear income taxation.[15]

This simplistic tax system is of interest for a variety of reasons. First, the use of linear tax instruments to finance public expenditures is conceivable in reality. Second, some models in the field of political economy are based on this assumption.[16] A normative model based on this tax structure allows the assessment of the welfare properties of the outcomes predicted by these studies. Finally, it turns out that the pattern of collective incentive problems that has been derived under a non-linear income tax is reversed under a linear tax on income. This proves the claim that the tax system itself is an essential determinant for an individual's assessment of public goods.

More precisely, the analysis is based on the assumption that individuals work less in response to an increased income tax rate. This assumption is shown to imply that individuals with a high level of earning ability suffer ceteris paribus from a larger utility loss if additional taxes are raised. Consequently the burden of taxation for a public good that is enjoyed by individuals of both classes is felt more intensively by the "rich".[17] This generates the following pattern of incentive problems: More able individuals tend to understate their willingness to pay for the public good because they suffer more intensively from an increase of the tax revenue requirement. Analogously, less able individuals exaggerate when asked about their valuation because they don't feel a large utility burden from higher taxes.

[15] A model in which a linear income tax is used to finance a public good and lump sum transfers as in Sheshinski (1972) and Hellwig (1986) would yield the same conclusions as the analysis in Chapter 4.

[16] See Polo (1998), Svensson (2000) or Persson and Tabellini (2000).

[17] Recall that the analysis of an optimal non-linear income tax in Chapter 4 yields the opposite effect.

As a consequence, collective incentive compatibility constraints require the use of excessive taxes, i.e. of taxes which are larger than actually needed to cover the cost of public good provision. Either they serve to make public good provision artificially expensive in order to prevent less able individuals from exaggerating their valuation of the public good; or, analogously, if more able individuals tend to understate their preferences, then excessive taxes are used to make non-provision of the public good less attractive. If these excessive taxes become very high, then an optimal provision rule does not incorporate all the information. Suppose for instance, that one needs to accompany public good provision with very high taxes in order to ensure that less able individuals reveal their valuation of the public good. Then, an optimal provision rule does not acquire information from them. Put differently, information that is too costly to obtain, is neglected by an optimal provision rule for public goods.

Chapter 2

Informative Voting and the Samuelson Rule

2.1 Introduction

We study a problem of optimal utilitarian public good provision in a continuum economy in which individuals have private information on their valuation of the public good and with uncertainty about the average valuation. An optimal rule reflects the average valuation. The higher this valuation, the higher should be the quantity that is provided. Consequently, an optimal provision rule relies on a procedure of information aggregation, i.e. prior to the final decision on public good provision, information on individual valuations has to be collected and to be aggregated.

Private information on public goods preferences gives rise to the classical *free-rider* problem in public good provision. This paper revisits the classical result of Arrow (1979) and d'Aspremont and Gérard-Varet (1979) that, in a quasi-linear economy with finitely many individuals, ex post efficiency, incentive compatibility and budget balance are simultaneously achievable; i.e. as long as individuals can be forced to contribute to the costs of public good provision the *free-rider* problem can be solved without a welfare burden, due to private information on public goods preferences.[1]

We argue that, in a limit economy with an infinite number of individuals, this result loses much of its appeal. In a finite economy with N individuals, incentive compatibility requires that individual contributions to the cost of public good provision are commensurate to the individual's impact on the quantity decision. In the limit case as $N \to \infty$, no single individual has a direct impact on public good provision. This implies that there is no possibility to make individual payment obligations dependent on announced preferences. Put differently, in the limit case, *equal cost sharing* is the only incentive compatible and feasible scheme of public goods finance. As a consequence, incentive requirements in the limit economy imply that there are *multiple equilibria*: all individuals pay equally for the public good and no single preference announcement has an effect on the chosen quantity, hence individuals are willing to make any conceivable announcement.

Our main concern, however, is not this multiplicity as such. Instead we ask what provision rules should be considered *"implementable"* in a large economy with uncertainty about the distribution of preferences. The multiplicity implies that any criterion of *"implementability"* in a large economy requires an assumption on how individuals break their indifference. Moreover, requiring only that individuals should be willing to reveal their preferences may yield an outcome that relies on the possibility to break individual indifference in a way that is in contrast with the individuals' interests concerning

[1] The more recent literature in this field has been concerned with the characterization of optimal allocations which are incentive compatible, budgetary feasible and, in addition, respect participation constraints. Recent contributions to this line of research are Hellwig (2003) or Norman (2004).

the outcome of the revelation game.

This point is most easily illustrated within the model that we investigate. Individuals have a quasi-linear utility function and possess private information on their *effective valuation* of the public good, which results from the interaction of a skill parameter and a taste parameter. Ceteris paribus, the effective valuation increases in the taste parameter and decreases in the skill parameter. The latter effect reflects that less skilled individuals suffer from a larger utility loss if forced to contribute to the cost of public good provision. The taste parameter is taken to be a binary variable; that is, individuals either have a low or a high taste parameter. Uncertainty about the average valuation of the public good results from the assumption that the percentage of individuals with a high taste realization is a random variable. An elicitation of the true state of the economy requires that individuals be asked about their actual taste parameters.

If, in this environments, one takes the view that *equal cost sharing* is a sufficient condition for *"implementability"*, then the optimal utilitarian rule for public good provision is a version of the Samuelson rule which equates, in every state of the economy, the *effective utilitarian valuation* of the public good and the marginal cost of public good provision. Now consider an individual with a high taste parameter but a very low skill level and, for the sake of concreteness, suppose this individual expects the state of the economy to be such that his own effective valuation of the public good is smaller than the effective utilitarian valuation. Consequently, under the Samuelson rule, this individual expects the level of the public good to be too large. If the individual takes his preferred outcome of the revelation game into account he should announce a low taste parameter and thereby *"contribute"* indirectly to a more preferred perception of the state of the economy. More generally, these considerations demonstrate that the Samuelson rule is implementable only if individuals reveal their preferences even though it would be in their interest to support a different outcome.

The aim of this paper is to propose a framework that takes account of these incentive issues and then to analyze the implications for the design of optimal mechanisms. To this end, we use an idea from the field of political economy, namely that voting mechanisms can be used for the purpose of information aggregation.

More precisely, the rule that we use to break individual indifference in the continuum economy is the following. We assume that an individual announces a high (low) taste parameter only if she prefers a rather large (small) level of the public good to be provided. Put differently, any announcement of a taste parameter is interpreted as a vote: a high taste parameter as a vote in favour of a large level of the public good and a low taste parameter as a vote in favour of a small level of the public good.

The assumption that individuals vote in favor of a large or small provision level only if this is their most preferred outcome of the revelation game translates into an additional incentive constraint. The average valuation of the public good can be inferred from the distribution of votes only if the public good is indeed provided in such a way that all individuals with a high taste parameter prefer a large provision level over a small level and all individuals with a low taste realization prefer a small level over a large level. We call these constraints the *informative voting (IV)* constraints.[2]

We characterize the optimal utilitarian provision rule which satisfies these

[2]These constraints resemble the notion of informative voting which is used in the field of political economy, see e.g. Austen-Smith and Banks (1996).

IV constraints. The main result is that the optimal extent of information aggregation is inversely related to a specific measure for the polarization of effective valuations of the public good in the economy.[3] To construct this measure, the economy is divided into two groups: those individuals with a high taste parameter and those individuals with a low taste parameter. Skill heterogeneity implies that there is *within group polarization* of effective valuations for the public good. E.g. an individual with low skills and a high taste parameter has a low effective valuation relative to an individual with high skills and a high taste parameter.

As soon as there is some degree of within-group polarization, an optimal provision rule exhibits *pooling*; i.e. the same provision level is chosen for a whole range of possible states of the economy. In an extreme case of within-group polarization, one finds that an optimal provision rule under IV constraints is such that the same quantity of the public good is provided in *every* state of the economy. In this sense, there is no use of information under an optimal mechanism.

Finally, to provide a theoretical foundation of the *informative voting* constraints we consider so called *sampling mechanisms*. A finite subset of N randomly drawn individuals is asked to report their taste parameters. Based on these taste announcements the mechanism designer, estimates the effective utilitarian valuation of the public good and decides on public good provision. Finally, the cost of provision is shared equally among all individuals in the economy.

We investigate the properties of an optimal mechanism as the sample size N

[3]These observations are similar in spirit to classical results from the signalling literature. See e.g. Crawford and Sobel (1982), Schultz (1996), Grossman and Helpman (2001, Ch. 4).

grows and individual influence on the level of public good provision disappears. We show that as $N \to \infty$ the optimal provision rule under sampling converges to the optimal provision rule in the original mechanism design problem under IV constraints. We interpret this result as providing a foundation for the IV requirement, which results from a simple, but somewhat arbitrary, rule for breaking indifference in a revelation game with a continuum of individuals.

As a corollary of this analysis, we show that the optimal provision rule under IV constraints provides an upper bound to the welfare levels which are achievable under sampling, for any finite sample size N. Put differently, we show that an optimal sample size does not exist. We interpret this observation as a version of the famous Condorcet Jury Theorem.[4]

The initial motivation of this paper was to study more generally, optimal rules for income taxation and public good provision in an economy where individuals have differing levels of ability and differing tastes for public goods.[5] Even though this paper focusses on the conceptual issues that arise in a large economy with private information on public goods preferences, it still provides a link between these two branches of the literature. It characterizes the optimal rule for public good provision in the following environment: Individuals derive utility from a public good, a private consumption good and leisure. Moreover, the utility function is additively separable and quasilinear

[4] A discussion of this theorem and of related results can be found in Piketty (1999).

[5] Heterogeneity with respect to earning abilities underlies the *equity-efficiency trade-off* studied in the theory of optimal income taxation in the tradition of Mirrlees (1971). Heterogeneity with respect to public goods preferences has been driving the literature on the *free-rider problem* in public goods provision (at least) since Clarke (1971) and Groves (1973).

in leisure.[6] The final allocation is determined sequentially. *First*, the level of public good provision is determined. This generates a revenue requirement in the public sector budget constraint. *Second*, the income tax schedule is chosen optimally subject to this predetermined revenue requirement.

The optimal provision for public goods rule derived in section 2.3 of this paper is also optimal in this extended model under the assumption that once the level of public good provision is fixed, tax authorities choose an optimal non-linear income tax in order to finance these expenditures.[7]

The remainder of this paper is organized as follows. In Section 5.2 we derive the mechanism design problem under IV constraints. In Section 2.3 the solution to this problem is characterized. Section 2.4 contains the discussion of sampling mechanisms and the derivation of the Condorcet Jury Theorem. The last section contains concluding remarks. All proofs can be found in the appendix.

[6] An optimal income tax in this setting has been characterized by Weymark (1986, 1987).

[7] In particular, tax authorities do not distort the optimal income tax in order to mitigate the welfare burden of the incentive constraints that are relevant for the decision on public good provision; that is, tax authorities cannot commit not to use an optimal income tax once the revenue requirement has been determined. However, as shown in Chapters 4 and 5, if such a commitment was possible it would, in general, lead to welfare superior outcomes.

2.2 The Model

2.2.1 The Environment

Individual Characteristics

The economy consists of a continuum of individuals $i \in I := [0,1]$. Individuals differ with respect to their skill level w^i, and their *valuation* of the public good, also referred to as their *taste* θ^i. The taste parameter may take two different values:

$$\theta^i \in \Theta := \{\theta_L, \theta_H\} \quad \text{with} \quad 0 \leq \theta_L < \theta_H,$$

where θ_L indicates a low taste for the public good and θ_H indicates a high taste. The skill parameter belongs to the compact interval

$$w^i \in W := [\underline{w}, \bar{w}] \quad \text{with} \quad 0 < \underline{w} \leq \bar{w} \ .$$

Individuals derive utility from the consumption of a public good, but don't like to contribute to the cost of provision. Agent i's utility function is given by

$$U^i \;\; = \;\; \theta^i Q - \frac{t^i}{w^i} \ .$$

Q denotes the quantity of a non-excludable public good and t^i captures i's contribution to the cost of public good provision. Note that a lower skill level implies a larger utility loss from a given payment obligation. The underlying idea is that, for less able individuals, it is harder to generate the income needed to meet a given payment obligation.

The function U^i is the cardinal representation of preferences which is relevant for welfare assessments. An individual's ranking of alternatives can be equivalently expressed by the monotone transformation

$$w^i U^i \;\; = \;\; \theta^i w^i Q - t^i \ .$$

We refer to the product $\theta^i w^i$ as individual i's *effective valuation* of the public good.

Informational Structure

The parameters w^i and θ^i are both private information of individual i and taken to be the realizations of the stochastically independent random variables $\tilde{w}^i$ and $\tilde{\theta}^i$, respectively. The random variables $\{\tilde{w}^i\}_{i \in I}$ are *independently and identically distributed (iid).* Their probability distribution is represented by a cumulative distribution function $F : W \to [0, 1]$ with density f. The random variables $\{\tilde{\theta}^i\}_{i \in I}$ are as well *iid.* We denote by p the individual probability of a high taste realization,

$$p \ := \ \text{Prob}\{\theta^i = \theta_H\} \ .$$

In addition, we assume that a *law of large numbers (LLN)* applies;[8] that is, almost surely, after the realization of randomness at the individual level, the cross-section distribution of characteristics in the economy coincides with the ex ante probability distribution that governs the randomness at the individual level. Accordingly, the value $F(w)$ and the probability p are interpreted as the fractions of individuals with earning ability $w^i \leq w$ and a high taste for the public good in the population, respectively. The *LLN* also implies that the empirical skill distribution and the empirical taste distribution are independent; that is, on every subinterval $[w', w''] \subset W$ of the support of the skill distribution, the share of individuals with a high taste realization is equal to p.

[8] Postulating a LLN for a continuum of *iid* random variables creates a measurability problem, as has been noted by Judd (1985) and Feldman and Gilles (1985). There is however a recent literature on modeling approaches which circumvent this measurability problem, see Alòs-Ferrer (2002) or Al-Najjar (2004).

We assume that the distribution F is common knowledge. Consequently, at the aggregate level, there is no uncertainty about the skill distribution. By contrast, the share of individuals with a high taste realization p is taken to be a random quantity; i.e. there is uncertainty with respect to the average valuation of the public good.

To sum up, the information structure has a known skill distribution and aggregate uncertainty with respect to the taste parameters. The unknown parameter p is henceforth also referred to as the *state of the economy*. It is the relevant object for the process of information aggregation.

Incentive Compatible Allocation Rules in a Continuum Economy

We limit attention to anonymous and incentive compatible allocation rules. An *anonymous allocation rule* (Q, t) consists of a *provision rule* for the public good and a *payment scheme* to cover the cost of provision.

The provision rule Q assigns to alternative values of p a quantity of the public good,

$$Q : [0, 1] \to \mathbb{R}_+, \quad p \mapsto Q(p).$$

This provision rule is anonymous in the sense that the level of provision $Q(p)$ depends only on the distribution of characteristics in the economy. I.e. it does not depend on the skill and taste realizations of specific individuals.[9]

The payment scheme t specifies for each individual i a payment obligation as a function of the distribution of characteristics in the economy p and individual i's characteristics (θ^i, w^i). The payment scheme is anonymous in the sense that individuals with the same characteristics have the same payment obligation, in every state p of the economy. Put differently, individual payments do not depend on the index i. Formally the payment scheme is

[9] Guesnerie (1995) calls this property *anonymity in influence*.

described as a function

$$t : [0,1] \times \Theta \times W \to \mathbb{R},\ (p, \theta, w) \mapsto t(p, \theta, w).$$

Individuals have private information on their skill and their taste parameter. This gives rise to the following incentive compatibility constraints.

Definition 2.1 An anonymous allocation rule is called incentive compatible (IC) if $\forall p \in [0,1]$, $\forall(\theta, w) \in \Theta \times W$, and $\forall(\hat{\theta}, \hat{w}) \in \Theta \times W$,

$$\theta w Q(p) - t(p, \theta, w) \quad \geq \quad \theta w Q(p) - t(p, \hat{\theta}, \hat{w}).$$

These incentive constraints are to be read as follows: Suppose that a mechanism designer wants to implement an allocation rule (Q, t). In a revelation game, he collects data from all individuals on their skill and on their taste parameter. The collection of these announcements is then used for two purposes: *first*, the profile of all taste announcements $(\hat{\theta}^i)_{i \in I}$ is used to deduce the actual value of p, *second*, for given p, the individual announcement $(\hat{\theta}^i, \hat{w}^i)$ is used to determine the payment obligation of individual i. The requirement of incentive compatibility deals with this *second* step only. It ensures that, for a given state p, an individual is indeed willing to make the payment prescribed by the payment scheme t. Put differently, an individual has no reason to hide the own characteristics in order to achieve a preferred treatment by the payment scheme.

These IC constraints have to be satisfied for each possible value of p. Put differently, whatever the *"announced state of the world"* which arises under the *first* step, any individual is willing to reveal the own characteristics truthfully. Hence, the underlying solution concept is one of *implementation in dominant strategies.*

As we consider a continuum economy, no single individual has a direct impact on the *"announced state of the world"*. This is reflected in the fact that the same level of p appears on the left hand side and the right hand side of the IC constraint. As a consequence, no single individual has a direct impact on the level of public good provision. Individuals are concerned only with a minimization of their payment obligation. This gives rise to the classical *free-rider problem*. As access to the public good is free, no one is willing to pay more than he is forced to.[10] These observations yield the following characterization of incentive compatible allocation rules.

Lemma 2.1 The following statements are equivalent.

1. (Q, t) is IC.

2. (Q, t) satisfies $\forall p$, $\forall(\theta, w)$ and $\forall(\hat{\theta}, \hat{w})$, $t(p, \theta, w) = t(p, \hat{\theta}, \hat{w})$.

Consequently, any IC payment scheme is constant, in the sense that, for given p, all individuals are treated equally. The converse statement is also true. That is, any anonymous provision rule $Q : p \mapsto Q(p)$ gives rise to an IC allocation rule if accompanied by constant payments, i.e. a payment scheme that does only depend on p.

Remark 2.1 An alternative characterization of incentive compatible allocation rules has been provided by Hammond (1979). He shows that an

[10] Note that if budget balance has to be achieved and there are limits to coercion due to participation constraints as in Mailath and Postlewaite (1990) or in Hellwig (2003) and there are individuals who do not value the public good at all – i.e. with effective valuation of 0 – one will end up with $Q \equiv 0$ under any admissible, incentive compatible allocation rule.

allocation rule is *IC* in the above sense, if and only if it is *decentralizable* by a suitably chosen tax policy. In the present setting, this tax system is, however, degenerate because it prescribes equal payments for all individuals. Still, at the conceptual level, this so called *taxation principle* has motivated the notion of incentive compatibility that we employ. It links our work to the field of public finance.

Budget Balance and Incentive Compatibility

The cost of public good provision is given by a twice continuously differentiable, strictly increasing and strictly convex cost function $K : \mathbb{R}_+ \to \mathbb{R}_+$, which satisfies $K(0) = 0$ as well as the boundary conditions

$$\lim_{x \to 0} K'(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} K'(x) = \infty.$$

The payment scheme has to be such that the costs of public good provision are covered, i.e. such that aggregate payments are equal to the cost of provision. Combining this requirement of budget balance with the requirement of *IC* yields the following observation.

Lemma 2.2 An anonymous allocation rule (Q, t) satisfies *IC* and budget balance if and only if the payment scheme is such that,

$$\forall p \in [0,1], \forall (\theta, w) \in \Theta \times W : \quad t(p, \theta, w) = K(Q(p)).$$

Lemma 2.2 allows to represent an individual's assessment of an allocation rule (Q, t), which is budgetary feasible and incentive compatible, in the following reduced form, which depends only on the provision rule Q,

$$U(p, \theta^i, w^i) \quad := \quad \theta^i Q(p) - \frac{K(Q(p))}{w^i} \ . \tag{2.1}$$

Expected utilitarian welfare

In what follows, we consider mechanism design problems of a benevolent utilitarian planner. A budgetary feasible and incentive compatible allocation is evaluated from an ex ante perspective, i.e. before the actual value of p is known. For simplicity, we assume that the planner takes the actual state of the economy p to be the realization of a random variable, $\tilde{p}$, which is uniformly distributed on the unit interval $[0, 1]$.[11]

Assumption 2.1 The mechanism designer takes p to be the realization of a random variable which is uniformly distributed on $[0, 1]$.

Under Assumption 2.1 and the *LLN*, ex ante expected utilitarian welfare becomes

$$\begin{aligned} EW &:= \int_0^1 \left\{ \left[p\theta_H + (1-p)\theta_L \right] Q(p) - \left[\int_{\underline{w}}^{\bar{w}} \frac{f(w)}{w} \, dw \right] K(Q(p)) \right\} dp \\ &= \lambda \int_0^1 \{ \bar{v}(p) Q(p) - K(Q(p)) \} \, dp \, , \end{aligned}$$

where $\lambda := \int (1/w) f(w) dw$ is an index of the marginal welfare effects of the cost of public good provision under equal cost sharing and

$$\bar{v}(p) := \frac{p\theta_H + (1-p)\theta_L}{\lambda}$$

is the *effective utilitarian valuation* of the public good.

2.2.2 The Problem of Information Aggregation

The problem of information aggregation is concerned with the question whether a mechanism designer is able to learn how many individuals in the economy

[11] Throughout we do not need to impose a common prior assumption. We only specify the prior beliefs of the mechanism designer.

have a high taste realization. Recall that the mechanism designer evaluates the profile of taste announcements $(\hat{\theta}^i)_{i\in I}$ to learn the actual value of p. Hence, he gets to know the actual state of the world if and only if (almost) all individuals reveal their taste parameter truthfully.

However, under an incentive compatible and budgetary feasible allocation rule individuals are indifferent which taste parameter to announce. I.e. the revelation game suffers from a problem of *multiple equilibria*. The reason is that we consider only anonymous allocation mechanisms. Consequently, no individual has a direct impact on public good provision. In addition, incentive compatibility requires that the payment scheme treats all individuals alike. These two facts imply that individuals are willing to make *any* announcement in the underlying revelation game.

In particular, this implies that individuals are willing to announce any taste parameter. The problem of information aggregation however is resolved only if all individuals announce their taste parameter truthfully. We will now argue that such an obedient behaviour cannot be taken for granted.

The problem with the Samuelson Rule

To illustrate this point we discuss the provision rule $Q^* : p \mapsto Q^*(p)$, which is chosen by a utilitarian planner who maximizes EW pointwise; i.e. who maximizes the expression $\bar{v}(p)Q(p) - K(Q(p))$ for every $p \in [0,1]$. This provision rule Q^* is nothing but the Samuelson rule under equal cost sharing. It is characterized by a continuum of first order conditions

$$\forall p : \quad \bar{v}(p) \quad = \quad K'(Q^*(p)) \; .$$

For brevity, we also refer to Q^* as the *first best provision rule.* Under Q^* individual preferences over the *"announced state of the world"* result from the reduced form

$$U^*(p,\theta,w) \quad := \quad \theta Q^*(p) - \frac{K(Q^*(p))}{w} \ .$$

It is easily verified that

$$U^*_p(p,\theta,w) \quad = \quad \frac{1}{w} Q^{*\prime}(p)\Big(\theta w - \bar{v}(p)\Big) \quad \begin{cases} < 0 & \text{if} \quad \theta w < \bar{v}(p) \ , \\ = 0 & \text{if} \quad \theta w = \bar{v}(p) \ , \\ > 0 & \text{if} \quad \theta w > \bar{v}(p) \ . \end{cases}$$

That is, under provision rule Q^* an individual prefers a larger level of p – or equivalently a larger level of public good provision – if and only if the own *effective valuation* exceeds the *effective utilitarian valuation*. Likewise an individual with an *effective valuation* below the average prefers to have a lower quantity of the public good.

These observations imply that an individual would refuse to reveal the own taste realization if he believed to have an influence on the decision on public good provision. To see this, consider an individual with a low taste realization and a high skill level which has an effective valuation close to $\theta_L \bar{w}$. Moreover, for the sake of concreteness, assume that this individual believes p to be very low.[12] If a vast majority of individual has a low taste realization, then this individual can be sure that its own effective valuation lies above the average, $\theta_L \bar{w} > \bar{v}(p)$. Put differently, under Q^*, the individual in question expects that the quantity of the public good is too low. As a consequence, the individual would be happy if the mechanism designer had a larger perception of p. Hence, this individual is inclined to announce a high taste realization in order to *"contribute"* to a more preferred outcome.

[12] I.e. when this individual decides ex interim what taste parameter to announce, her prior beliefs put a lot of probability mass on values of p which are close to zero.

The Informative Voting Constraints

The point of these considerations is that, even though individuals have no direct influence on public good provision, they are not indifferent regarding the mechanism designer's perception of p. That is, they are not indifferent regarding the outcome of the revelation game.

We now state a formal condition, called *informative voting (IV)*, that we impose on the mechanism design problem. It is inspired by game-theoretic models of voting decisions in the field of political economy. For the moment, we just introduce these conditions and discuss their interpretation. However, in Section 2.4 we discuss a more rigorous theoretical foundation.

In our setting IV ensures that individuals are *"really"* willing to reveal their taste parameter. That is, the IV constraints guarantee that individuals are not tempted to break the indifference among all conceivable taste announcements such that they *"contribute"* with a false announcement to a more preferred state perception.

Definition 2.2 A provision rule Q is said to satisfy the IV property if the following holds for any $w \in W$ and any $p \in [0,1]$: $U(p, \theta_L, w)$ is non-increasing in p and $U(p, \theta_H, w)$ is non-decreasing in p.

Referring to these monotonicity postulates as IV constraints is based on the idea that any individual subscribes to one of two groups, either to those individuals with $\theta^i = \theta_L$ or to the group with $\theta^i = \theta_H$. Informative voting hence is a sufficient condition which ensures that each individual supports the group which shares the own taste characteristic. As a consequence of this behavior, the distribution of votes allows to deduce the actual value of p.

We interpret the IV constraints as a condition of *robustness*;[13] that is, they ensure that, whatever the prior beliefs of individuals on the likelihood of different values of p, no individual has a reason to report a false taste parameter in order to *"contribute"* to a more favorable state perception.[14] This is most clearly seen, if Q is a differentiable function of p. In this case, the IV constraints become: for all p and for all w,

$$U_p(p, \theta_L, w) \leq 0 \quad \text{and} \quad U_p(p, \theta_H, w) \geq 0 \; .$$

If the provision rule Q satisfies these IV constraints, then, for all realizations of p, an individual with a low taste realization prefers to live in an economy, where less individuals have a high taste realization. Likewise, all individuals with a high taste realization are better off if p is larger.

Mechanism Design under IV constraints

We can now define the mechanism design problem of a utilitarian planner who has to choose an IC allocation rule (Q, t) and in addition uses the IV constraints to ensure that he can deduce the actual value of p from the profile of taste announcements $(\hat{\theta}^i)_{i \in I}$ in the revelation game.

Definition 2.3 The following problem is called the *informative voting problem (P)*: Choose a provision rule Q in order to maximize EW subject to the

[13] Further discussion of these notions of incentive compatibility can be found in Bergemann and Morris (2005), Chung and Ely (2004) or Kalai (2004). For mechanism design problems with private values, the notions *robustness* and *implementation in dominant strategies* are equivalent.

[14] In particular, even in the extreme case in which the true value of p is known to all individuals and the mechanism designer is the only uninformed party, all individuals are willing to reveal their taste parameter. This case gives rise to a *mechanism design problem under complete information.* See Moore (1992) for a survey.

IV constraints. The solution to this problem is denoted by Q^{**}, the induced optimal welfare level by EW^{**}.

In Section 2.4 we provide a theoretical foundation of this mechanism design problem under IV constraints. There, we show that the optimal provision rule Q^{**} can be interpreted as the limit outcome of a sequence of mechanism design problems with vanishing individual influence on public good provision. Before turning to this issue, we characterize the solution of problem P in the next section.

2.3 Optimal Provision under Informative Voting

In this section we characterize the solution to the *informative voting problem*. The key insight is that the extent by which the optimal provision rule Q^{**} reflects variations in the average valuation of the public good p depends on a specific measure of preference polarization. The role of skill heterogeneity for preference polarization is easily demonstrated with the following alternative characterization of the IV property.

Lemma 2.3 A provision rule Q satisfies IV if and only if the following two properties hold for any pair p, p' with $p' > p$:

i) Q is increasing: $Q(p') \geq Q(p)$.

ii) If $Q(p') > Q(p)$, then

$$\theta_H \underline{w} \;\; \geq \;\; \frac{K(Q(p')) - K(Q(p))}{Q(p') - Q(p)} \;\; \geq \;\; \theta_L \bar{w} \; .$$

The Lemma follows from standard arguments. It says that IV is equivalent to the requirements that a provision rule must be monotonically increasing in p and, moreover, that an individual with effective valuation $\theta_L \bar{w}$ always prefers a small provision level over a large provision level, whereas an individual with effective valuation $\theta_H \underline{w}$ always prefers the large provision level.

Hence, to satisfy the IV constraints only the preferences of the extreme types with effective valuations $\theta_L \bar{w}$ and $\theta_H \underline{w}$, respectively, have to be taken into account. Intuitively, if even an individual with the top skill level prefers a small quantity of the public good over a larger quantity in case of a low taste realization, then the same is true for any individual with $\theta^i = \theta_L$ and an effective valuation $\theta^i w^i \leq \theta_L \bar{w}$. Consequently, under a monotonic provision rule, any individual with a low taste parameter wants the perceived state to be as small as possible and this implies that the individual's IV constraint is satisfied. Analogously, for individuals with a high taste parameter only the IV condition for an individual with the minimal skill level has to be taken into account. These observations allow to prove the following proposition.

Proposition 2.1

i) If $\theta_L \bar{w} > \theta_H \underline{w}$, then a provision rule Q satisfies the IV property if and only if it is constant: for all p, $Q(p) = x$, for some $x \in \mathbb{R}_+$.

ii) The first best provision rule Q^* satisfies the IV constraints if and only if there is no skill heterogeneity, i.e. $\underline{w} = \bar{w}$.

Part i) of Proposition 1 shows that the requirement of IV may indeed heavily restrict the set of admissible provision rules. If $\theta_L \bar{w} > \theta_H \underline{w}$, then a provision rule satisfies IV if and only if information aggregation plays no role. The underlying logic is as follows. If an individual with effective valuation $\theta_L \bar{w}$

prefers provision level $Q(p)$ over the larger provision level $Q(p')$, then the same is true for an individual with an effective valuation smaller than $\theta_L \bar{w}$. Hence, if $\theta_L \bar{w} > \theta_H \underline{w}$ there exist individuals with a high taste realization who prefer $Q(p)$ as well. But IV requires these individuals to prefer $Q(p')$. As a consequence, all these statements are consistent with each other only if $Q(p') = Q(p)$.

We interpret such a parameter constellation with $\theta_L \bar{w} > \theta_H \underline{w}$ as one of *large within-group polarization.* This terminology reflects the following considerations: The IV constraints essentially require that all individuals with the same taste realization have the same views on public good provision. However, there is *within-group polarization* of preferences due to skill heterogeneity. We take the distance $\bar{w} - \underline{w}$ to be a measure of *within-group polarization.* It is said to be *large* if there are individuals in the low taste group who have a skill level which is so high that their effective valuation exceeds the one of low skilled individuals in the high taste group.

A parameter constellation with $\underline{w} = \bar{w}$ is one in which there is *no within-group polarization* at all. This implies that the IV constraints do not have any bite and that the first best provision rule Q^* is admissible. To see this note that without *within-group polarization*, there are only two possible effective valuations, a high one and a low one. The utilitarian planner cares about the average. If he decides on public good provision, without taking the IV constraints into account, then, whatever the actual state of the economy, the chosen provision level will be too high for individuals with a low valuation; that is, those individuals would be happy if the planner believed that the fraction of individuals with a low taste realization was in fact larger. Hence, their IV constraint is satisfied. The same is true for individuals with a high taste realization.

This reasoning does not go through if there is some degree of *within-group polarization*. As has been shown in the previous section, under Q^* there exist values of p such that individuals with a low taste parameter but a very high skill level would be happy if the planner believed that the share of individuals with a high taste realization was larger.

The results in Proposition 2.1 characterize the optimal provision rule under IV constraints for the extreme cases of *no within-group polarization* and *large within-group polarization*. For the remainder of this section we consider parameter constellations of *moderate within-group polarization* which satisfy $\underline{w} \neq \bar{w}$ and $\theta_L \bar{w} \leq \theta_H \underline{w}$. As follows from part ii) of Proposition 2.1, in these cases the first best provision rule Q^* is not available. The question thus becomes what an optimal deviation from Q^* looks like. In the following we first provide a taxonomy of possible solutions to the informative voting problem under *moderate within-group polarization* and identify three relevant classes of provision rules. In a second step, we argue that the degree of *within-group polarization* determines the class to which the actual solution belongs.

As has been established in Lemma 2.3, to satisfy the IV constraints only the extreme types with effective valuations $\theta_L \bar{w}$ and $\theta_H \underline{w}$ matter.

An implication of this observation is that a provision rule with the IV property has *at most* one provision level that falls short of the most preferred provision level of an individual with effective valuation $\theta_L \bar{w}$, henceforth denoted by $\bar{Q}_L$ and formally defined by the condition

$$\{\bar{Q}_L\} \quad = \quad \text{argmax}_Q \; \theta_L \bar{w} Q - K(Q) \; .$$

To see this suppose, to the contrary, that there are two provision levels below $\bar{Q}_L$. Then, due to the fact that the function $\theta_L \bar{w} Q - K(Q)$ is single-peaked, an individual with effective valuation $\theta_L \bar{w}$ prefers the larger of these two. But IV rules out this possibility. The analog reasoning establishes that there can

be *at most* one provision level that exceeds the most preferred provision level of an individual with effective valuation $\theta_H \underline{w}$, denoted by $\underline{Q}_H$.
Under *moderate within-group polarization* one has

$$Q^*(0) \quad < \quad \bar{Q}_L \quad < \quad \underline{Q}_H \quad < \quad Q^*(1) \; .$$

Hence, a utilitarian planner would want to choose a continuum of different provision levels smaller than $\bar{Q}_L$ and as well a continuum of different provision levels larger than $\underline{Q}_H$. The *IV* constraints imply, however, that he can choose at most one such provision level.
It is shown below that an optimal provision rule has *exactly* one provision level below $\bar{Q}_L$ and *exactly* one exceeding $\underline{Q}_H$. To describe the relevant provision rules with this property some additional terminology is needed.

Definition 2.4 An increasing provision rule $Q_4 : p \mapsto Q_4(p)$ is said to have *four pooling levels* if

$$Q_4(p) \quad := \quad \begin{cases} Q_4^s & \text{for} \quad 0 \leq p \leq \hat{p} \; , \\ \hat{Q}_4^s & \text{for} \quad \hat{p} < p < \hat{p}' \; , \\ Q^*(p) & \text{for} \quad \hat{p}' \leq p \leq \tilde{p}' \; , \\ \tilde{Q}_4^l & \text{for} \quad \tilde{p}' < p < \tilde{p} \; , \\ Q_4^l & \text{for} \quad \tilde{p} \leq p \leq 1 \; , \end{cases}$$

where Q_4^s and $\hat{Q}_4^s$ satisfy $\theta_L \bar{w} \; Q_4^s - K(Q_s^4) = \theta_L \bar{w} \; \hat{Q}_4^s - K(\hat{Q}_4^s)$, i.e. an individual with effective valuation $\theta_L \bar{w}$ is indifferent between these two provision levels. Likewise, $\tilde{Q}_4^l$ and Q_4^l satisfy $\theta_H \underline{w} \tilde{Q}_4^l - K(\tilde{Q}_4^l) = \theta_H \underline{w} Q_4^l - K(Q_4^l)$.

Finally, the critical indices are implicitly defined by the equations:[15]

$$\bar{v}(\hat{p}) = \theta_L \bar{w} \ , \quad Q^*(\hat{p}') = \hat{Q}_4^s \ , \quad Q^*(\tilde{p}') = \tilde{Q}_4^l \ , \quad \bar{v}(\tilde{p}) = \theta_H \underline{w} \ .$$

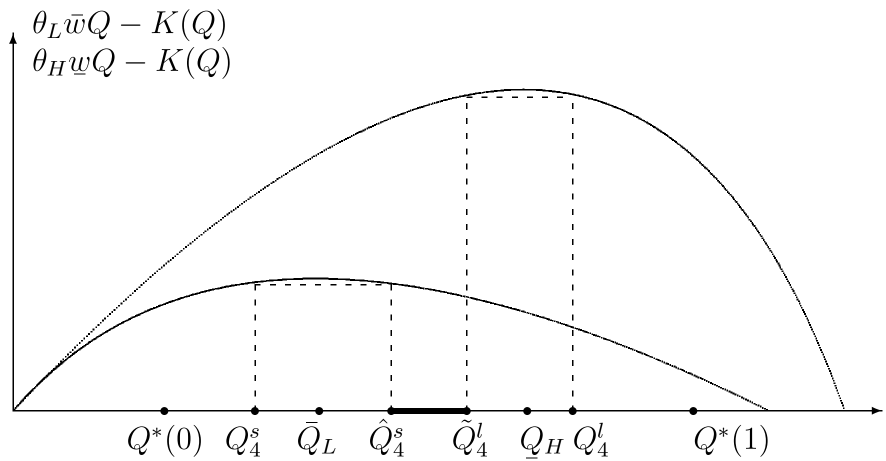

Figure 1: The figure depicts a provision rule characterized by four pooling levels Q_4^s, $\hat{Q}_4^s$, $\tilde{Q}_4^l$ and Q_4^l. Over an intermediate range the provision level equals $Q^*(p)$.

Provision rules with four pooling levels are such that Q_4^s is linked with $\hat{Q}_4^s$ via a binding IV constraint for an individual with effective valuation $\theta_L \bar{w}$. Likewise, Q_4^l is linked with $\tilde{Q}_4^l$ via an IV constraint for an individual with $\theta_H \underline{w}$. Moreover, there is a range of p for which the provision level is equal to $Q^*(p)$, the provision level that would be chosen in the absence of IV-constraints. As a consequence, if the smallest pooling level Q_4^s is close to $Q^*(0)$ and the largest pooling level Q_4^l is close to $Q^*(1)$, then a provision

[15] This already presumes an optimal choice of the critical indices. To see this, note that a utilitarian planner will choose e.g. $\hat{p}$ according to the following criterion: Let $Q(p) = Q_4^s$ if and only if $\bar{v}(p)Q_4^s - K(Q_4^s)$ exceeds $\bar{v}(p)\hat{Q}_4^s - K(\hat{Q}_4^s)$. Given the binding IV constraint which links Q_4^s and $\hat{Q}_4^s$, this is equivalent to $Q(p) = Q_4^s$ if and only if $\bar{v}(p) \leq \theta_L \bar{w}$.

rule with four pooling levels approximates provision rule Q^*. A degenerate case of a provision rule with four pooling levels arises if the range where the provision rules Q_4 and Q^* coincide shrinks to a singleton. We say that such a provision rule is characterized by *three pooling levels*. It has a medium sized provision level Q_3^m that is linked via a binding IV constraint with a small pooling level Q_3^s and via a binding IV constraint with a large pooling level Q_3^l. Finally, a provision rule with no provision level between $\bar{Q}_L$ and $\underline{Q}_H$ is characterized by *two pooling levels* .

Definition 2.5 An increasing provision rule $Q_2 : p \mapsto Q_2(p)$ is said to have *two pooling levels* if,

$$Q_2(p) \quad := \quad \begin{cases} Q_2^s & \text{for} \quad 0 \leq p \leq \bar{p} \, , \\ Q_2^l & \text{for} \quad \bar{p} < p \leq 1 \, , \end{cases}$$

where $Q_2^s \leq \bar{Q}_L$, $Q_2^l \geq \underline{Q}_H$ and $\bar{p}$ is defined implicitly be the equation

$$\bar{v}(\bar{p})Q_2^s - K(Q_2^s) \quad = \quad \bar{v}(\bar{p})Q_2^l - K(Q_2^l) \; .$$

Proposition 2.2 Suppose there is skill heterogeneity ($\underline{w} \neq \bar{w}$). A provision rule which solves the informative voting problem either makes no use of information or has two, three or four pooling levels.

This Proposition shows that a provision rule whose image lies entirely between $\bar{Q}_L$ and $\underline{Q}_H$ cannot be optimal. In the appendix it is shown by a Lagrangean approach that one can always find a provision rule with four pooling levels that is superior to such a truncated provision rule. The same argument can be used to show that an optimal provision rule has exactly one element smaller than $\bar{Q}_L$ and exactly one element larger than $\underline{Q}_H$; that is, also partial truncations can be excluded. These considerations single out the above candidates.

This taxonomy allows to solve problem P in the following way. One has to compare the welfare levels that can be realized with a constant provision rule, a provision rule with *two pooling levels* , a provision rule with *three pooling levels* and a provision rule with *four pooling levels.* Generally, this requires to solve for each class a separate optimization problem and to rank the resulting welfare levels.

There is however a general intuition, to which class the optimal provision rule belongs, depending on the parameters of the model. Reconsider Figure 1 and note that if $\bar{Q}_L$ is close to $Q^*(0)$ and $\underline{Q}_H$ is close to $Q^*(1)$, then a provision rule with four pooling levels is close to Q^*, which is optimal if IV is not required. This suggests that if the *within-group* polarization of views on the optimal level of public good provision is relatively mild – in the sense that all individuals with taste parameter θ_L want to have a provision level in a neighborhood of $Q^*(0)$ and all individuals with θ_H want to have a provision level similar to $Q^*(1)$ – then one ends up with a provision rule which exhibits four pooling levels and approximates the first best rule Q^*.

However, if the difference between $\bar{Q}_L$ and $\underline{Q}_H$ – or equivalently the difference between $\theta_H \underline{w}$ and $\theta_L \bar{w}$ – shrinks, so does the range over which a provision rule with four pooling levels coincides with Q^*. There will be a critical parameter constellation such that the monotonicity constraint $\hat{Q}_4^s \leq \tilde{Q}_4^l$ binds and one ends up with three pooling levels .

As the *within-group* polarization increases further, the difference $\underline{Q}_H - \bar{Q}_L$ becomes very small. Then a provision rule with three pooling levels needs to have all three provision levels very close to each other. Hence, there is only very little use of information as a provision rule with three pooling levels becomes similar to one with $Q(p) = const$, for all p. In such a case a provision rule with only two pooling levels, which are however to a larger extent differentiated from each other, is superior; that is, a provision rule with only two pooling levels eventually becomes more attractive.

We refrain from providing a general proof of these intuitive statements. This would require an awkward exercise, which distinguishes a variety of assumptions on the parameters $\theta_L, \theta_H, \bar{w}, \underline{w}$ and λ, i.e. the skill distribution F. We only provide an example which allows to verify the intuition developed above.

Example Suppose $K(Q) = \frac{1}{2}Q^2$, $\theta_L = 1$, $\theta_H = 3$, and $\lambda = 1$. Let $\underline{w} = 1 - x$ and $\bar{w} = 1 + x$. In this example x is a measure of the welfare burden imposed by the requirement of IV. This welfare burden vanishes as $x \rightarrow 0$ implying that $\underline{w} \rightarrow \bar{w}$. As $x \rightarrow \frac{1}{2}$ one converges to the case with $\theta_L \bar{w} = \theta_H \underline{w}$ which precludes any information aggregation. One may verify that for sufficiently small x, a provision rule with four pooling levels is optimal. For $x \geq 2^{-\frac{3}{2}}$, an optimal provision rule with four pooling levels is transformed into the degenerate case with only three pooling levels . Finally, for x close to $\frac{1}{2}$ a provision rule with only two pooling levels is superior.

2.4 Sampling

In section 2.2 we observed that in a continuum economy, the problem of information aggregation has no structure because individuals are indifferent which taste parameter to announce. For the definition of the *informative voting problem* we just assumed that individuals break this indifference based on their most preferred state perception. The purpose of this section is to derive the IV constraints in a way that avoids this ad-hocery.

We discuss informative voting decisions by a finite random sample of N individuals. In a finite sample, each sample member has a strictly positive influence on the mechanism designer's state perception. This structure can be used to study the limit case as individual influence gets arbitrary small. We thus regard the sampling approach, as a way to single out the *"reasonable"* outcome in a continuum economy. Indeed, as we will show below, as $N \to \infty$, the *optimal provision rule based on sampling* converges to the provision rule which solves the *informative voting problem P*. We view this result as the ultimate foundation of the idea that *informative voting* is the relevant constraint for the problem of information aggregation.

More precisely, we analyze the following mechanism design problem: Individual preferences are given by the reduced form utility function $U(p, \theta, w)$. A mechanism designer tries to learn the actual value of p. To this end he draws a random sample of N individuals and asks those individual to report a low or a high valuation of the public good, or, equivalently, asks the sampled individuals to vote. Based on these N preference announcements, the mechanism designer forms *beliefs* about the actual state of the world p. The final decision on public good provision is a function of those beliefs and hence dependent on the preference announcements of the sampled individuals, or, equivalently, on the distribution of votes in the sample. As a consequence, there is a need of appropriate incentives for sampled individuals: in a finite sample, sampled individuals have a strictly positive impact on the mechanism designer's beliefs. He will thus learn the true sample distribution of characteristics, only if he decides on public good provision in such a way that indeed each sampled individual is willing to reveal the own taste realization truthfully. Put differently, information aggregation is possible only if sample members are willing to vote informatively.

Remark 2.2 The sampling procedure that we study is based on the reduced form representation of individual utility. Sampled individuals thus internalize the consequences of their announcements for a *given* scheme of public goods finance. As a consequence our approach differs from the sampling mechanisms analyzed by Green and Laffont (1979, Ch.12) and Gary-Bobo and Jaaidane (2000). While these authors study as well allocation problems where only a subset of individuals is used for information aggregation, they assume that contributions to the cost of public good provision may differ for individuals within the sample and those who are not in the sample. By contrast, we assume that there is a *tax system* that treats all individuals identically, irrespective of whether or not they are sample members.

Mechanism design based on sampling

For the purpose of information aggregation, the mechanism designer communicates with a random sample S_N of N individuals. He uses the number $m = \#\{i \in S_N \mid \theta^i = \theta_H\}$ of high taste realizations to update his prior beliefs on the actual state p of the economy. These updated beliefs give rise to a posterior density function ϕ_N. I.e. the density ϕ_N formalizes the notion of the mechanism designer's *perceived state of the economy.*

Lemma 2.4 Suppose that Assumption 2.1 holds and that there are m high taste realizations in a sample of size N. The conditional density $\phi_N(\cdot \mid m)$ is given by

$$\phi_N(p \mid m) \quad = \quad (N+1)\tbinom{N}{m}p^m(1-p)^{N-m}\ . \tag{2.2}$$

Based on the state perception $\phi_N(\cdot \mid m)$ the mechanism designer decides on public good provision. That is, he chooses a *provision rule based on sampling* *of size* N,

$$Q_N : \{0, 1, \ldots, N\} \rightarrow \mathbb{R}_+, \quad m \mapsto Q_N(m) \ .$$

A scheme of public goods finance has to satisfy *incentive compatibility* and *budget balance.* As a consequence, *equal cost sharing* is the only admissible payment scheme and preferences over the level of public good provision can be represented in reduced form,

$$U(m, \theta, w) \quad = \quad \theta Q_N(m) - \frac{K(Q_N(m))}{w} \ .$$

In a revelation game, each sampled individual has an impact on the number m of high taste realizations which are observed by the mechanism designer. The following incentive conditions ensure that each sampled individual is willing to reveal the own taste realization truthfully. We call those constraints the *informative voting under sampling of size* N (IV_N) constraints.

Definition 2.6 A provision rule Q_N allows for *informative voting under sampling of size N (IV_N)* if the following inequalities hold for all $m \in \{0, \ldots, N-1\}$ and for all $w \in W$:

$$\theta_L w Q_N(m) - K(Q_N(m)) \quad \geq \quad \theta_L w Q_N(m+1) - K(Q_N(m+1)) \ ,$$

$$\theta_H w Q_N(m) - K(Q_N(m)) \quad \leq \quad \theta_H w Q_N(m+1) - K(Q_N(m+1)) \ .$$

The IV_N constraints ensure that the truth is a dominant strategy in a revelation game, in which individuals announce either a high or a low taste parameter and preferences are given in reduced form. Put differently, IV_N achieves *robustness* in the sense that *ex post* no sample member would want to revise his taste announcement in order to improve the quantity of the public good installed under provision rule Q_N.

We assume that the mechanism designer chooses Q_N in order to maximize utilitarian welfare from the *ex ante* perspective. Ex ante the sample distribution m is unknown. The mechanism designer takes m to be the realization of a random variable which behaves in accordance with the planner's prior beliefs and the Law of Large Numbers. As shown in the appendix, this implies that the mechanism designer takes m to be the realization of a random variable which is uniformly distributed over the support $\{0, 1, \ldots, N\}$. Consequently, an explicit expression for this utilitarian objective function can be derived.

Lemma 2.5 Under Assumption 2.1, a provision rule based on sampling of size N, Q_N, gives rise to the following level of expected utilitarian welfare

$$EW_N \quad := \quad \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v} \left(\frac{m+1}{N+2} \right) Q_N(m) - K(Q_N(m)) \right\} .$$

According to Lemma 2.5, a mechanism designer who observes a sample in which m individuals have a high valuation of the public good ends up with an effective valuation of the public good given by

$$\bar{v} \left(\frac{m+1}{N+2} \right) \quad = \quad \frac{m+1}{N+2} \frac{\theta_H}{\lambda} + \frac{N-m+1}{N+2} \frac{\theta_L}{\lambda} .$$

Note that this effective valuation is strictly increasing in m, and for all $m \in \{0, \ldots, N\}$ it exceeds $\bar{v}(0)$ and falls short of $\bar{v}(1)$.

Definition 2.7 The following problem is called the *finite sample problem* P_N: Choose a provision rule based on sampling of size N, Q_N, in order to maximize EW_N subject to the IV_N constraints. The solution to this problem is denoted by Q_N^{**}, the induced optimal welfare level by EW_N^{**}.

In the following we will study the behavior of Q_N^{**} and EW_N^{**} as $N \to \infty$. Before, we state an alternative characterization of the IV_N requirement which is entirely analogous to the characterization of the IV property in Lemma 2.3.

Lemma 2.6 A provision rule based on sampling Q_N satisfies IV_N, if and only if the following two properties hold for any pair $m, m' \in \{0, 1, \ldots, N\}$ with $m' > m$:

i) Q_N is increasing: $Q_N(m') \geq Q_N(m)$.

ii) If $Q_N(m') > Q_N(m)$, then

$$\theta_H \underline{w} \quad \geq \quad \frac{K(Q_N(m')) - K(Q_N(m))}{Q_N(m') - Q_N(m)} \quad \geq \quad \theta_L \bar{w}\ .$$

Large sample properties

We will now derive the main result of this section, namely that as $N \to \infty$ the *informative voting problem* P and the *finite sample problem* P_N are essentially equivalent.

We start with the observation, that the maximal welfare level EW^{**}, which is achievable under IV constraints, is, for any sample size N, an upper bound for the expected welfare which is achievable under IV_N constraints. To establish this claim, we define, for any given $N \in \mathbb{N}$, the following piecewise constant continuation of the solution of problem P_N, which we denote by $\tilde{Q}_N^{**}$:

$$\tilde{Q}_N^{**} : [0,1] \quad \to \quad \{Q_N^{**}(m)\}_{m=0}^{N} \quad \text{with}$$

$$\begin{aligned} \tilde{Q}_N^{**}(p) \quad &:= \quad Q_N^{**}(m) \quad \text{for} \quad \frac{m}{N+1} \leq p < \frac{m+1}{N+1}, \\ \tilde{Q}_N^{**}(1) \quad &:= \quad Q_N^{**}(N). \end{aligned} \tag{2.3}$$

The welfare level induced by $\tilde{Q}_N^{**}$ is denoted $\widetilde{EW}_N^{**}$.

Lemma 2.7 For any $N \in \mathbb{N}$, the following inequalities hold:

$$EW_N^{**} \;\leq\; \widetilde{EW}_N^{**} \;\leq\; EW^{**} \, .$$

The first inequality is strict if and only if Q_N^{**} is not constant.

The inequality $\widetilde{EW}_N^{**} \leq EW^{**}$ is obvious because the provision rule $\tilde{Q}_N^{**}$ satisfies IV and hence is an admissible choice for *problem P*. The remainder of the proof then is to verify that the construction works. There is no deep insight to be gained from this exercise. Lemma 2.7 is the key in order to derive the main results of this section.

Proposition 2.3 As $N \to \infty$, the welfare level which is realized under a solution of the *finite sample problem* P_N, EW_N^{**}, converges to the welfare level which is realized under a solution of the *informative voting problem P*, EW^{**}. Formally,

$$\lim_{N\to\infty} EW_N^{**} \;=\; EW^{**} \, .$$

The proof is based on the following construction. Start out from the provision rule Q^{**} which solves *problem P* and define its *restriction* $Q_{|N}^{**}$ to the domain of m as follows: for each $m \in \{0, 1, \ldots, N\}$, let $Q_{|N}^{**}(m) := Q^{**}(m/N)$. Using that Q^{**} satisfies the IV constraints, one easily verifies that $Q_{|N}^{**}$ has the IV_N property. This implies that the welfare level $EW_{|N}^{**}$ which results from $Q_{|N}^{**}$ has to be smaller than the one which results from a solution to *problem* P_N, i.e. $EW_{|N}^{**} \leq EW_N^{**}$. However as $N \to \infty$ the difference between $Q_{|N}^{**}$ and Q^{**} vanishes, i.e. $EW_{|N}^{**} \to EW^{**}$. Combining these observations with Lemma 2.7 implies that Proposition 2.3 must be true.

The Proposition basically shows that as $N \to \infty$ the concepts of *informative voting under sampling* and *informative voting* are equivalent in terms of their welfare implications. If there is a *unique* solution Q^{**} of *problem* P, this equivalence is strengthened: the optimal provision rule under IV constraints, Q^{**}, and the optimal provision rule under IV_N constraints Q_N^{**} *"coincide"* in the limit.

Corollary 2.1 Suppose there is a unique solution Q^{**} to problem P, and let $\tilde{Q}_N^{**}$ be as defined in (2.3). Then, for all $p \in [0,1]$

$$\lim_{N\to\infty} \tilde{Q}_N^{**}(p) \quad = \quad Q^{**}(p) \ .$$

A Condorcet Jury Theorem

As a byproduct of the preceding analysis we can prove a version of the famous *Condorcet Jury Theorem*. This theorem is concerned with decision making in committees. In its most simple version,[16] the theorem says that whenever each committee member has some private information on the state of the world and, moreover, all committee members have identical preferences, then a larger committee size is always preferable. The underlying logic, is that a larger committee has more pieces of information available and will thus undertake the "right" decision with a larger probability.

For our version of the *Condorcet Jury Theorem* we interpret the random sample S_N as a committee.[17] We then ask the question whether there is a finite optimal sample size. We will show that, whenever some information aggregation is desirable, then, for any N there exists $N' > N$ such that

[16]See Austen-Smith and Banks (1996) and Piketty (1999). A more advanced treatment can be found in Feddersen and Pesendorfer (1997).

[17]A similar approach can be found in Auriol and Gary-Bobo (2005).

$EW_{N'} > EW_N$.

It has been shown in Lemma 2.7 that $EW_N^{**} \leq EW^{**}$, i.e. a planner who evaluates a continuum of taste reports but is constrained by the requirement of IV will never do worse than a planner who just uses the reports of a finite sample of individuals under IV_N constraints. As stated in the following Proposition, whenever a solution to the *informative voting problem* P is not degenerate, this inequality is strict.

Lemma 2.8 Suppose that provision rule Q^{**}, which solves P, is not constant.[18] Then for any $N \in \mathbb{N}$, the following inequality holds: $EW_N^{**} < EW^{**}$.

Combining this observation with Proposition 2.3 yields the desired result.

Corollary 2.2 Suppose that provision rule Q^{**} is not constant. Then for any given $N \in \mathbb{N}$ there exists $N' \in \mathbb{N}$ with $N' > N$ such that $EW_N^{**} < EW_{N'}^{**}$.

Whenever some degree of information aggregation is desirable, there is no optimal sample size. The intuition for this result is the following. A growing sample size N implies that the mechanism designer's estimate of the actual state of the economy becomes more precise. This allows for a better adjustment of the final provision level to the actual state of the economy. However, a larger N also implies a larger set of IV_N constraints. These additional constraints, however, do not undermine this reasoning. A mechanism designer with a large sample can always mimic a small sample outcome by choosing not to use certain pieces of information. Hence, larger samples generate additional degrees of freedom for the mechanism designer.

[18]Sufficient conditions are: $\theta_L \bar{w} \leq \theta_H \underline{w}$, $\hat{Q}^*(0) \leq Q^*(1)$ and $Q^*(0) \leq \tilde{Q}^*(1)$.

2.5 Concluding Remarks

We have addressed a problem of public goods provision in a continuum economy with private information of individuals on their valuation of a public good and uncertainty about the average valuation. As has been shown, the requirement of an incentive compatible payment scheme gives rise to a problem of multiple equilibria in the underlying revelation game.

We have formulated two different approaches to deal with this problem. The first, rather naive, idea is a simple criterion for breaking individual indifference: whenever an individual is literally indifferent among all conceivable announcements in a revelation game, use the individual's preferences over the composition of the economy to break this indifference. I.e. whenever an individual is indifferent between, say, announcements a and b but would be happy if more individuals in the economy announced b, then assume that the individual in question will announce b as well.[19]

The second approach, *informative voting under sampling*, distinguishes more explicitly between *information aggregation* to determine the optimal quantity of a public good and the *financing* of this desired quantity. A large random sample of individuals is used for the process of information aggregation. Sampled individuals now have an impact on public good provision and this governs their behavior in the revelation game. Hence, the multiple equilibrium problem is eliminated.

The crucial assumption is that the payment scheme treats sampled individuals not differently as compared to individuals who possess the same charac-

[19]In models with voting over two alternatives and a continuum of voters one often finds the statement that this behavior is the only one which survives the elimination of weakly dominated strategies. Implicitly, this reasoning appeals to a large but finite economy. Examples include Gersbach (2005) or Meirowitz (2005).

teristics but have not been in the sample. From a general mechanism design perspective, this assumption clearly involves a loss of generality. There certainly exist welfare superior allocation mechanisms, which do not share this property. Hence, it has to be emphasized, that we ask a more special question, namely how a scheme of *taxation*, which treats all individuals equally for public goods finance, should be designed under a need of information aggregation.

Finally we have shown, that, for large random samples, these two different approaches, are equivalent. That is, the simple rule which we refer to as *informative voting* can be interpreted as the limit outcome of vanishing individual influence under a voting mechanism with a finite number of individuals.

A third approach which also provides a foundation of the *informative voting* constraints can be found in Chapter 3. That paper allows agents to form coalitions in order to manipulate the mechanism designer's perception of the state of the world. An admissible provision rule for public goods then has to fulfill a condition which eliminates incentives for manipulative collective actions. It is shown in Chapter 3 that for the simple quasi-linear environment analyzed in this paper, to achieve *coalition-proofness* it suffices to prevent the formation of coalitions which are arbitrary small but have strictly positive mass. This requirement is then shown to be equivalent to the *informative voting* constraints.

The common feature of Chapter 3 and the present chapter is, that in order to get a foundation of incentive constraints in the continuum, one has to grant individuals some small influence on public good provision. This can be achieved either by considering their impact in a large, but finite, random sample, or by considering the scope for collective action in small neighborhoods with positive mass.

2.6 Appendix

Proof of Lemma 2.3. We first show that the IV constraints imply statements i) and ii). The IV constraints imply that the following two inequalities have to hold,

$$\theta_L w Q(p) - K(Q(p)) \ \ \geq \ \ \theta_L w Q(p') - K(Q(p')) \ ,$$

$$\theta_H w Q(p) - K(Q(p)) \ \ \leq \ \ \theta_H w Q(p') - K(Q(p')) \ .$$

Adding up these inequalities yields:

$$(\theta_H - \theta_L) w [Q(p') - Q(p)] \ \ \geq \ \ 0 \ .$$

This establishes i). Now suppose that $Q(p) < Q(p')$. Then for any $w \in W$, IV requires that

$$\frac{K(Q(p')) - K(Q(p))}{Q(p') - Q(p)} \ \ \geq \ \ \theta_L w \ .$$

This property holds for all $w \in W$ if and only if it holds for the largest skill level $\bar{w}$,

$$\frac{K(Q(p')) - K(Q(p))}{Q(p') - Q(p)} \ \ \geq \ \ \theta_L \bar{w} \ .$$

Likewise we derive the requirement

$$\theta_H \underline{w} \ \ \geq \ \ \frac{K(Q(p')) - K(Q(p))}{Q(p') - Q(p)} \ .$$

This establishes ii).

The proof that i) and ii) imply that the IV property holds is now immediate.

■

Proof of Proposition 2.1. Statement i) is a direct consequence of Lemma 2.3. It thus remains to be shown that Q^* satisfies IV if and only if $\underline{w} = \bar{w}$. To prove the "only if part", suppose that $\underline{w} \neq \bar{w}$. Consider the indirect utility function $U^*(p, \theta, w)$ As shown in the body of the text U^* is increasing in p as long as $\theta w > \bar{v}(p)$, i.e. the individual's effective valuation of the public good exceeds the effective utilitarian valuation. Analogously, U^* is decreasing in p if θw falls short of the utilitarian valuation. Now consider a level of p such that[20]

$$\theta_L \bar{w} \quad > \quad \bar{v}(p) \quad > \quad \frac{\theta_L}{\lambda}.$$

This implies that there exists a critical value $\hat{w} \in]\underline{w}, \bar{w}[$ such that all individuals with $\theta^i = \theta_L$ and $w^i > \hat{w}$ have an effective valuation $\theta_L w^i$ exceeding $\bar{v}(p)$. Therefore, they would prefer a slightly larger perceived value of p. This violates the IV property.

To prove the "if part", suppose that $\underline{w} = \bar{w} =: \tilde{w}$. As Q^* is a strictly increasing function, Lemma 2.3 implies that Q^* satisfies IV if and only if $p' > p$ implies that

$$\theta_H \tilde{w} \quad \geq \quad \frac{K(Q^*(p')) - K(Q^*(p))}{Q^*(p') - Q^*(p)} \quad \geq \quad \theta_L \tilde{w}\ .$$

We show in the following that the convexity of K and the first order conditions characterizing Q^* imply that these inequalities are indeed satisfied for any pair p' and p with $p' > p$. From the convexity of the cost function we have

$$K'(Q^*(p')) \quad > \quad \frac{K(Q^*(p')) - K(Q^*(p))}{Q^*(p') - Q^*(p)} \quad > \quad K'(Q^*(p))\ .$$

[20] As $\bar{v}(p)$ is a convex combination of $\frac{\theta_H}{\lambda}$ and $\frac{\theta_L}{\lambda}$, for any $x \in [\frac{\theta_L}{\lambda}, \frac{\theta_H}{\lambda}]$ there exists p such that $\bar{v}(p) = x$.

With $\underline{w} = \bar{w} =: \tilde{w}$, the first order conditions characterizing Q^* imply

$$K'(Q^*(p)) \quad = \quad \bar{v}(p) \quad = \quad \tilde{w}(p\theta_H + (1-p)\theta_L) \quad \geq \quad \tilde{w}\theta_L \;,$$

$$K'(Q^*(p')) \quad = \quad \bar{v}(p') \quad = \quad \tilde{w}(p'\theta_H + (1-p')\theta_L) \quad \leq \quad \tilde{w}\theta_H \;.$$

■

Proof of Proposition 2.2. The required arguments are lengthy but not difficult. Hence the arguments are only sketched. As has been shown in Proposition 2.1, there are parameter constellations such that Q^{**} is indeed constant. Now suppose that Q^{**} is not constant. Thus, by Proposition 2.1, $\theta_L\bar{w} \leq \theta_H\underline{w}$ and hence $Q^*(0) \leq \bar{Q}_L \leq \underline{Q}_H \leq Q^*(1)$. We show within the next three steps that Q^{**} is a provision rule with either two, three or four pooling levels .

Step 1. Denote by $\mathcal{V}_Q$ the image of a provision rule Q, i.e. $x \in \mathcal{V}_Q$ if and only if there exists $p \in [0,1]$ with $Q(p) = x$. Under the IV constraints, there exists at most one element $x \in \mathcal{V}_Q$ with $x < \bar{Q}_L$.

Proof of step 1. To see this, suppose to the contrary that there exist $x, y \in \mathcal{V}_Q$ with $x < y < \bar{Q}_L$. Under IV, as characterized in Lemma 2.3, this implies that there exist p and $p' > p$ with $Q(p) < Q(p') < \bar{Q}_L$. This yields

$$\theta_L\bar{w}Q(p) - K(Q(p)) < \theta_L\bar{w}Q(p') - K(Q(p')) \;,$$

a contradiction to the IV requirement for an individual with effective valuation $\theta_L\bar{w}$. Analogously one shows that the image of an admissible provision rule contains at most one element x with $x > \underline{Q}_H$.

Step 2. We now show that a provision rule Q for which there exists $y \in \mathcal{V}_Q$ with $y \in [\bar{Q}_L, \underline{Q}_H]$ is a candidate for a solution only if there exist as well $x, z \in \mathcal{V}_Q$ with $x < \bar{Q}_L$ and $\underline{Q}_H < z$.

Proof of step 2. To this end, we first argue that a provision rule Q for which there exist neither $x \in \mathcal{V}_Q$ with $x < \bar{Q}_L$ nor $z \in \mathcal{V}_Q$ with $z > \underline{Q}_H$ cannot be optimal. Such a hypothetical provision rule would satisfy $\mathcal{V}_Q \subset [\bar{Q}_L, \underline{Q}_H]$. But this, for such a provision rule to be *optimal*, would imply even $\mathcal{V}_Q = [\bar{Q}_L, \underline{Q}_H]$. However, this would be the degenerate case of a provision rule with four pooling levels, which results as the limit outcome as Q_4^s converges to $\bar{Q}_L$ and Q_4^l converges to $\underline{Q}_H$. Under a provision rule characterized by four pooling levels expected welfare EW satisfies the following equation:

$$\begin{aligned}\frac{EW}{\lambda} = \; & \hat{p}\left[\bar{v}\left(\frac{\hat{p}}{2}\right)Q_4^s - K(Q_4^s)\right] + (\hat{p}' - \hat{p})\left[\bar{v}\left(\frac{\hat{p}' + \hat{p}}{2}\right)\hat{Q}_4^s - K(\hat{Q}_4^s)\right] \\ & + \int_{\hat{p}'}^{\tilde{p}'}\left\{\bar{v}(p)Q^*(p) - K(Q^*(p))\right\}dp + (\tilde{p} - \tilde{p}')\left[\bar{v}\left(\frac{\tilde{p}' + \tilde{p}}{2}\right)\tilde{Q}_4^l - K(\tilde{Q}_4^l)\right] \\ & + (1 - \tilde{p})\left[\bar{v}\left(\frac{1 + \tilde{p}}{2}\right)Q_4^l - K(Q_4^l)\right] ,\end{aligned}$$

where $\hat{Q}_4^s$ and $\hat{p}'$ are implicit functions of Q_4^s. Similarly, $\tilde{Q}_4^l$ and $\tilde{p}'$ are implicit functions of Q_4^l. Taking these functional relationships into account one may compute the partial derivatives and verify that

$$\lim_{Q_4^s \to \bar{Q}_L} \frac{\partial EW(Q_4^s, Q_4^l)}{\partial Q_4^s} < 0 \quad \text{and} \quad \lim_{Q_4^l \to \underline{Q}_H} \frac{\partial EW(Q_4^s, Q_4^l)}{\partial Q_4^l} > 0.$$

Thus, $Q_4^s = \bar{Q}_L$ and $Q_4^l = \underline{Q}_H$ cannot be optimal.

We now argue in a similar manner that it cannot be optimal to choose a provision rule such that there exist $y, z \in \mathcal{V}_Q$ with $\bar{Q}_L < y < \underline{Q}_H < z$, but such that there does not exist $x \in \mathcal{V}_Q$ with $x < \bar{Q}_L$.

Define $\tilde{z} < \underline{Q}_H$ by the equation $\theta_H \underline{w}\, z - K(z) = \theta_H \underline{w}\, \tilde{z} - K(\tilde{z})$. Note that for such a provision rule to be a optimal under IV it has to be true that $y \leq \tilde{z}$ and that $\mathcal{V}_Q = [\bar{Q}_L, \tilde{z}] \cup \{z\}$ by optimality and step 1. Again, this is a degenerate case of a provision rule with four pooling levels , namely the one that results as Q_4^s converges to $\bar{Q}_L$ and $Q_4^l = z$. As above this hypothetical solution can be ruled out as

$$\lim_{Q_4^s \to \bar{Q}_L} \frac{\partial EW(Q_4^s, Q_4^l)}{\partial Q_4^s} \quad < \quad 0 \, .$$

The analogous argument allows to rule out a provision rule such that there exist $x, y \in \mathcal{V}_Q$ with $x < \bar{Q}_L < y < \underline{Q}_H$ but such that there does not exist $z \in \mathcal{V}_Q$ with $z > \underline{Q}_H$.

Step 3. We now claim that a provision rule, for which there exist $x, y \in \mathcal{V}_Q$ with $\bar{Q}_L < x < y < \underline{Q}_H$, is a candidate for a solution only if the whole interval satisfies $[x, y] \subset \mathcal{V}_Q$.

Proof of Step 3. By step 2, there are $a, b \in \mathcal{V}_Q$ with $a < \bar{Q}_L < \underline{Q}_H < b$. Define $\tilde{b} < \underline{Q}_H$ by the equation $\theta_H \underline{w}\ b - K(b) = \theta_H \underline{w}\ \tilde{b} - K(\tilde{b})$. Analogously, define $\hat{a} > \bar{Q}_L$ by $\theta_L \bar{w}\ a - K(a) = \theta_L \bar{w}\ \hat{a} - K(\hat{a})$. For the hypothesized provision rule to be a optimal under IV it has to be true that, $\hat{a} \leq x < y \leq \tilde{b}$ and that $[x, y] \subset [\hat{a}, \tilde{b}] \subset \mathcal{V}_Q$.

Steps 1-3 imply that an optimal provision rule under IV which is not constant has to be one with two, three or four pooling levels.

■

Proof of Lemma 2.4. The mechanism designer's prior beliefs are given by the density function ϕ. Under Assumption 2.1, $\phi(p) = 1$ for all $p \in [0, 1]$. Let ν be the number of agents with high taste parameters in a sample of size N. Again, from an ex ante perspective ν is a random variable. If one uses repeatedly that

$$\int_0^1 p^m (1-p)^{N-m} dp \quad = \quad \frac{m!(N-m)!}{(N+1)!}$$

one can verify the following statement,

$$\begin{aligned} pr(\nu = m) &= \int_0^1 pr(\nu = m \mid p)\phi(p)dp \\ &= \int_0^1 \binom{N}{m} p^m (1-p)^{N-m} dp \quad = \quad \frac{1}{N+1}\ . \end{aligned} \tag{2.4}$$

This is intuitive, with p uniformly distributed, all possible realizations of ν are equally likely. Now suppose that $\nu = m$ and consider the conditional density ϕ_N thereby induced over p. By Bayes' rule

$$\phi_N(p \mid \nu = m) \;=\; \frac{pr(\nu = m \mid p)\phi(p)}{pr(\nu = m)} = (N+1)\binom{N}{m} p^m (1-p)^{N-m}\ .$$

■

Proof of Lemma 2.5. At the *interim* stage, after observing m, the mechanism designer updates his beliefs on p. Expected welfare at the interim stage is hence given by

$$\begin{aligned} EW_N^{int}(m) = \;& \lambda \mathrm{E}[\bar{v}(p)Q_N(m) - K(Q_N(m)) \mid m] \\ =\;& \lambda \int_0^1 [\bar{v}(p)Q_N(m) - K(Q_N(m))]\phi(p \mid \nu = m)dp \\ =\;& \lambda (N+1)\binom{N}{m}\Big(\int_0^1 \Big[\frac{p\theta_H + (1-p)\theta_L}{\lambda}\, Q_N(m) - K(Q_N(m))\Big] \times \\ & \quad p^m(1-p)^{N-m}\, dp\Big) \\ =\;& \lambda \left[\frac{m+1}{N+2}\frac{\theta_H}{\lambda} + \frac{N-m+1}{N+2}\frac{\theta_L}{\lambda}\right] Q_N(m) - K(Q_N(m)). \end{aligned}$$

From the *ex-ante* perspective, the outcome m of the sampling procedure is the realization of a random variable, which we denote by ν. Taking expectations over m, using (2.4), expected welfare from the *ex-ante* perspective equals

$$
\begin{aligned}
EW_N = & \sum_{m=0}^{N} EW^{int}(m)pr(\nu = m) \\
= & \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m+1}{N+2}\right) Q_N(m) - K(Q_N(m)) \right\}.
\end{aligned}
$$

■

Proof of Lemma 2.7. By definition, $\tilde{Q}_N^{**}$ is monotonically increasing in p and inherits the IV property from the fact that $\{Q_N^{**}(m)\}_{m=0}^{N}$ is RS_N. This is obvious from the characterization of RS_N in Lemma 2.6 and the characterization of IV in Lemma 2.3. Hence, by the optimality of Q^{**} among the provision rules satisfying IV, $\widetilde{EW}_N^{**} \leq EW^{**}$. It thus remains to be shown that $EW_N^{**} \leq \widetilde{EW}_N^{**}$.

In order to compute $\widetilde{EW}_N^{**}$, we first collect a number of observations which are easily verified by the reader.

1. For all $\underline{p}, \bar{p} \in [0,1]$
$$\int_{\underline{p}}^{\bar{p}} \bar{v}(p)dp = (\bar{p} - \underline{p})\bar{v}\left(\frac{\bar{p} + \underline{p}}{2}\right).$$
2. For all $m \in \{0, 1, \ldots, N\}$
$$\frac{m + \frac{1}{2}}{N+1} = \frac{m+1}{N+2} + \frac{m - \frac{1}{2}N}{(N+1)(N+2)}.$$
3. For all $x, y \in [0,1]$ with $x + y \in [0,1]$, $\bar{v}(x+y) = \bar{v}(x) + \frac{\theta_H - \theta_L}{\lambda} y$.
4. By definition of EW_N^{**} and Q_N^{**},
$$EW_N^{**} = \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m+1}{N+2}\right) Q_N^{**}(m) - K(Q_N^{**}(m)) \right\}.$$

Using these equalities one arrives at

$$\begin{aligned}\widetilde{EW}^{**}_N &= \lambda\int_0^1 \left\{\bar{v}(p)\tilde{Q}^{**}_N(p) - K(\tilde{Q}^{**}_N(p))\right\} dp \\ &= EW^{**}_N + \frac{\theta_H - \theta_L}{(N+1)^2(N+2)}\sum_{m=0}^{N}(m - \frac{1}{2}N)Q^{**}_N(m)\ .\end{aligned}$$

To complete the proof we show that $\sum\limits_{m=0}^{N}(m-\frac{1}{2}N)Q^{**}_N(m) \geq 0$. This expression equals

$$\textstyle\sum_{m=0}^{\frac{1}{2}N}(\frac{1}{2}N - m)(Q^{**}_N(N-m) - Q^{**}_N(m))$$

if N is even and

$$\textstyle\sum_{m=0}^{\frac{N-1}{2}}(\frac{1}{2}N - m)(Q^{**}_N(N-m) - Q^{**}_N(m))$$

if N is odd. However, as Q^{**}_N is increasing, those sums are non-negative. Moreover, they are strictly positive, and hence $EW^{**}_N < \tilde{EW}^{**}_N$, if and only if Q^{**}_N is not constant.

■

Proof of Proposition 2.3. Let Q^{**} be a solution of problem P and $Q^{**}_{|N}$ its *restriction* to the domain $\{0, 1, \ldots, N\}$, Formally, for each $m \in \{0, 1, \ldots, N\}$, $Q^{**}_{|N}(m)$ is defined by the equation

$$Q^{**}_{|N}(m) \;:=\; Q^{**}\left(\frac{m}{N}\right)\ .$$

Using that Q^{**} satisfies the IV constraints, one easily verifies that $Q^{**}_{|N}$ has the IV_N property.

Denote by $EW^{**}_{|N}$ the expected welfare level induced by $Q^{**}_{|N}$. Then, since Q^{**}_N is optimal among the provision rules with the IV_N property, $EW^{**}_{|N} \leq EW^{**}_N$.

Moreover,

$$\begin{aligned} EW^{**}_{|N} &= \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m+1}{N+2}\right) Q^{**}\left(\frac{m}{N}\right) - K(Q^{**}\left(\frac{m}{N}\right)) \right\} \\ &= \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m}{N}\right) Q^{**}\left(\frac{m}{N}\right) - K(Q^{**}\left(\frac{m}{N}\right)) \right\} \\ &\quad + \frac{\theta_H - \theta_L}{N(N+1)(N+2)} \sum_{m=0}^{N} (N-2m) Q^{**}\left(\frac{m}{N}\right). \end{aligned}$$

The first term in this sum is a so-called *Riemann sum*[21] for $\bar{v}(p)Q^{**}(p) - K(Q^{**}(p)$ and thus converges to EW^{**} for growing N. The second term in the sum is bounded from above by the expression $\frac{\theta_H - \theta_L}{N+2} Q^{**}(1)$, which vanishes as $N \to \infty$. Consequently,

$$\lim_{N\to\infty} EW^{**}_{|N} = EW^{**}.$$

Summing up and using Lemma 2.7, the following chain of inequalities must hold:

$$EW^{**} = \lim_{N\to\infty} EW^{**}_{|N} \leq \lim_{N\to\infty} EW^{**}_{N} \leq EW^{**}.$$

■

Proof of Corollary 2.1. By Lemma 2.7 and Proposition 2.3,

$$\lim_{N\to\infty} EW^{**}_{N} = \lim_{N\to\infty} \widetilde{EW}^{**}_{N} = EW^{**}.$$

$\tilde{Q}^{**}_N$ is IV_N for all $N \in \mathbb{N}$. Thus, the uniqueness of Q^{**} among the provision rules which satisfy IV and yield welfare level EW^{**} implies the claimed property of pointwise convergence.

■

[21] See e.g. Heuser (1998, Ch.10).

Proof of Lemma 2.8. Denote by EW^u the maximal level of expected welfare which can be generated by some constant provision rule. If the solution to problem P is not constant, then $EW^u < EW^{**}$. If, for $N \in \mathbb{N}$, a solution to problem P_N is constant, then $EW_N^{**} = EW^u < EW^{**}$. If, by contrast, Q_N^{**} is not constant, then $EW_N^{**} < EW^{**}$ by Lemma 2.7.

■

Chapter 3

Collectively Incentive Compatible Tax Systems

3.1 Introduction

This chapter introduces the concept of a *collectively incentive compatible tax system* as a tool that allows the study of two incentive problems simultaneously. The first incentive problem stems from the fact that individuals have private information on their earning ability. This restricts the set of admissible tax systems in a way analyzed in the theory of optimal income taxation in the tradition of Mirrlees (1971). The second incentive problem arises because of the fact that individuals have private information on their valuation of a non-excludable public good. This yields the classical free-rider problem in public good provision. Individuals like to enjoy the public good but are not willing to pay for it. This restricts the set of admissible provision rules for public goods.

The joint treatment of these two incentive problems fills a gap in the theory of public economics. This gap exists because the normative theory of public goods provision has two separate branches. On the one hand, there is the *theory of optimal taxation*. This theory assumes that there is a large economy and that the tax setting institution can be assumed to know the distribution of characteristics in the economy. This institution has some set of instruments available which include various direct and indirect tax instruments, as well as the quantities in which public goods are provided. Given some welfare assessment, the tax setting institution solves for an optimal scheme of taxation and public goods provision under a public sector budget constraint. The optimal quantity of a public good is then determined according to some modified version of the classical *Samuelson rule*, named after Samuelson (1954), which takes the use of distortionary tax instruments to finance public good provision into account.[1]

In this approach there is no problem of information aggregation. The economy is large. This justifies the assumption that the distribution of characteristics is taken to be commonly known. Consequently, there is no need to elicit individual valuations of public goods.

The second branch of the literature on public goods provision is driven by this latter problem. I refer to it as the *mechanism design* approach.[2] In the simplest setting, a benevolent mechanism designer has to choose a provision rule for a public good and a payment scheme. He wants to provide a public good in such a way that the level of provision reflects the average valuation

[1] Examples of this approach include Atkinson and Stern (1974), Wilson (1991), Boadway and Keen (1993), Nava et al. (1996), Sandmo (1998), Hellwig (2005b, 2004) and Gaube (2000, 2005).

[2] This literature originates from the study of Vickrey-Clarke-Groves Mechanisms, see Clarke (1971) and Groves (1973). A survey can be found in Laffont (1987) or the textbook of Mas-Colell et al. (1995). Recent contributions to this line of research are Hellwig (2003) and Norman (2004).

of those individuals who enjoy the public good; i.e. the higher the average valuation, the more of the public good should be provided. Obviously, such a provision rule requires that the average valuation of the public good be learned. Accordingly, the mechanism is used for the purpose of information aggregation.

Such an analysis is typically undertaken for an economy consisting of finitely many individuals. This implies that every single individual's valuation is an important quantity for a determination of the average valuation. Consequently, each single is able to influence on the level of public good provision and hence the enjoyment of the public good by all other individuals. Public good provision in a finite economy thus becomes a rather complex strategic game, driven by multilateral externalities.

To summarize this brief overview, the idea that a reasonable criterion for public good provision requires the collection of information on valuations of the public good and that this causes an incentive problem has been addressed in finite economies but not in large economies and not in conjunction with the tax instruments which are used to finance public expenditures. The present paper provides a framework that allows these issues to be addressed.

At a conceptual level, this raises the question of an appropriate solution concept. The main issue is whether information aggregation is really an incentive problem in a large economy. To see this, suppose there are infinitely many individuals, each with private information about his or her own valuation of the public good. Moreover, assume that the average valuation of the public good is not known to the institution that decides on public good provision. This institution uses a revelation game to learn this average valuation; i.e. it collects data from all individuals and computes the average.

Based on this exercise, it determines the quantity of a public good.

In a large economy no single individual has a direct impact on the average valuation of the public good, as perceived by the institution in charge of public good provision. Hence, no single individual has a direct reason to hide his true valuation of the public good. Viewing the problem from this perspective, which focusses on individual incentives, would lead to the conclusion that, in a large economy, information aggregation does not involve an incentive problem.

The present paper, however, takes a different view. It is assumed that individuals can form coalitions in order to manipulate jointly the perceived average valuation of the public good and hence the decision on public good provision. Consequently, the institution in charge of public good provision learns the true average valuation only if there is no large group of individuals that benefits from a collective manipulation of the announced profile public goods preferences. Allocations which do not provoke such strategic manipulations by groups of agents are henceforth called *collectively incentive compatible*.

The formalism developed below introduces this idea into the setup typically used in the literature on optimal income taxation; i.e. individuals have private information on their earning ability. Simultaneously it is assumed that this uncertainty about individual productivity levels disappears in the aggregate and that the cross-section distribution of earning ability is commonly known. The link between income taxes and public goods arises via a public sector budget constraint. It is required that tax revenues are sufficient to cover the cost of public good provision.

The new assumptions introduced in this paper are that, in addition to the private information on earning ability, individuals have private information on their valuation of a public good. Moreover, this uncertainty about individual valuations does not wash out in the aggregate. There is aggregate uncertainty because the joint cross-section distribution of earning ability and valuations of the public good is not commonly known.

The characterization of the set of implementable allocation is treated as a problem of mechanism design in a large economy.[3] An allocation consists of an income tax schedule and a provision rule for public goods. To be implementable it has to fulfill three requirements. First, it has to be *feasible*. Second, it has to be *individually incentive compatible (I-IC)*: From a single individual's perspective there is no reason to hide one's characteristics, taking the announcements of all other individuals in the revelation game as given. Finally, it has to be *collectively incentive compatible (C-IC)*: No coalition of individuals has an incentive to engage in a collective manipulation of the decision on public good provision, taking the behavior of individuals outside the coalition as given.

The main formal result of the paper provides a characterization of individually and collectively incentive compatible tax systems that is useful in applications. If preferences of individuals are additively separable between private and public goods, then one can separate *individual* and *collective* incentive problems. *Individual incentives* deal with a screening problem, namely of identifying individual levels of earning ability. *Collective incentives* address the problem of information aggregation that arises because the joint distribution of earning ability and public goods preferences is not commonly known. The separability result shows that *collective incentive compatibility* holds if

[3]This approach has been introduced by Hammond (1979) and Guesnerie (1995). See Hellwig (2004) for a recent contribution.

no coalition of individuals benefits from a manipulation of public goods preferences, taking as given that these individuals reveal their earning ability. Put differently, there is no need to worry about coalitions that manipulate the announced profile of earning ability.

While this is per se not a deep insight, it proves convenient for a more explicit characterization of implementable allocations in more specific environments. To illustrate this, one such application is studied in more detail, namely an economy in which individuals have quasi-linear preferences over the quantity of a public good and their individual payment obligation.[4]

Moreover, the separability result is used in Chapter 4 and in Chapter 5.[5] These papers study an economy with only two groups of individuals, those with a high and those with a low level of earning ability. Chapter 4 shows that, in order to ensure *collective incentive compatibility*, it suffices to exclude collective manipulations which are such that all individuals with the same level of earning ability jointly misreport their valuation of the public good.

The remainder of the paper is organized as follows. Section 3.2 contains the formal description of the economy. In addition, the example of a quasi-linear economy is used to demonstrate that an optimal scheme of income taxation and public good provision is in general vulnerable to the formation of manipulating coalitions. In section 3.3 the solution concept of a *collectively incentive compatible tax system* is introduced. This section also

[4]The analysis uses some results from Chapter 2. That paper studies the same environment but is concerned with voting as a mechanism that solves the problem of information aggregation.

[5]These papers differ in the set of available tax instruments. In Chapter 5 the proceeds from a linear tax on income are used to finance public good provision. Chapter 4 studies the case of optimal nonlinear income taxation.

contains a discussion of the related literature on mechanism design problems under coalition formation. Section 3.4 derives the result that, with separable preferences, a separation of *individual* and *collective* incentive problems is possible. In Section 3.5 this observation is used to characterize the optimal *I-IC* and *C-IC* allocation in the quasi-linear economy. The last section contains concluding remarks. All proofs are in the appendix.

3.2 The Problem

3.2.1 The environment

There is a large set of individuals identified with the unit interval $I = [0,1]$ and equipped with measure μ. An individual $i \in I$ has a utility function U defined over the quantity $Q \in \mathbb{R}_+$ of a non-excludable public good, and bundles of private goods $A \in \mathbb{R}^l$. In addition, utility depends on individual characteristics. I distinguish a taste parameter $\theta_i \in \Theta$, $\Theta \subset \mathbb{R}_+$, to formalize heterogeneity regarding valuations of the public good and a productivity or skill parameter $w_i \in W$, $W \subset \mathbb{R}_+$. For brevity, I denote a pair of individual characteristics (θ_i, w_i) by γ_i and the set $\Theta \times W$ by Γ. U is thus written as

$$U = U(Q, A, \gamma_i) \, .$$

Example 3.1 In the theory of optimal income taxation A is a pair (C, Y) consisting of consumption of private goods $C \in \mathbb{R}_+$ and effective labour supply or income $Y \in \mathbb{R}_+$. In this setting, the productivity parameter captures individual heterogeneity with respect to the utility loss associated with a given level of effective labour supply.

When discussing applications, I impose the assumption that the utility function U is additively separable in the utility contribution of the public good, depending on the taste parameter θ_i, and the utility contribution of A, depending on the skill parameter w_i.

Assumption 3.1 The utility function U is additively separable:

$$U \quad = \quad v(Q, \theta_i) + u(A, w_i) \ .$$

The assignment of characteristics to individuals is represented by an assignment function $\gamma_a : I \to \Gamma$ with image denoted by $\{\gamma_i\}_{i \in I} = \{(\theta_i, w_i)\}_{i \in I}$. It is assumed throughout that there is *assignment uncertainty*. I.e. the function γ_a – or equivalently the profile $\{\gamma_i\}_{i \in I}$ – is not commonly known. Instead, individual i has private information on the parameter γ_i.

Assumption 3.2 Almost all assignments γ_a are measurable functions.

Assumption 3.2 implies that expressions such as e.g.

$$\mu(\{i \mid \theta_i \leq \theta \text{ and } w_i \leq w\}) \quad \text{or} \quad \mu(\{i \mid \theta_i \leq \theta \ \})$$

are, for any resolution of assignment uncertainty γ_a, well defined.

In addition to assignment uncertainty, there is *aggregate uncertainty* referring to the empirical distribution of individual characteristics in the economy. From an ex ante perspective there are different states of the economy. Each such state corresponds to a cross-section distribution of characteristics and is represented by a cumulative distribution function (cdf) $D : \Gamma \to [0, 1]$ that lists for each $\gamma = (\theta, w)$ the fraction of individuals with characteristics $\gamma_i \leq \gamma$,

$$D(\gamma) = \mu(\{i \mid \theta_i \leq \theta \text{ and } w_i \leq w\}) \ .$$

Assumption 3.3 There is aggregate uncertainty, in the sense that the actual cross-section distribution of characteristics D in the economy is not commonly known. There is a commonly known set $\mathcal{D}$ of possible states of the economy.

The following information about the distribution of characteristics in the economy is common knowledge. There is *aggregate stability* regarding the marginal distribution of productivity parameters: any feasible distribution $D \in \mathcal{D}$ gives rise to the same marginal cumulative distribution function F, with $F(w) = \mu(\{i \mid w_i \leq w\})$, of the skill parameter in the economy.

Assumption 3.4 There is aggregate stability with respect to the productivity parameter; i.e. for any $D \in \mathcal{D}$ there is a commonly known marginal cumulative distribution function $F : W \to [0, 1]$.

Remark 3.1 At this general level, there is no need to be more specific on the relation between randomness at the individual level – i.e. the precise nature of assignment uncertainty – and Assumptions 3.3 and 3.4 on the aggregate structure of the economy. In the literature one often finds that $\{\gamma_i\}_{i \in I}$ is taken to be the realization of stochastic process consisting of independent and identically distributed *(iid)* random variables. In addition, with appeal to some *Law of Large Numbers for Large Economies*, any realization of this process is assumed to induce an assignment that is consistent with the assumptions imposed on the aggregate features of the economy. A mathematical foundation for this approach is provided by Al-Najjar (2004).

The following example is used repeatedly to illustrate the main ideas of this paper.

3.2.2 An Example

Let $W = [\underline{w}, \bar{w}]$ be a compact interval and let F be such that there exists a density f that is strictly positive on $(\underline{w}, \bar{w})$. Let $\Theta = \{\theta_L, \theta_H\}$ with $\theta_L < \theta_H$. *Aggregate uncertainty* is formalized as follows: Denote by p the fraction of individuals with a high taste parameter, $p = \mu(\{i \mid \theta_i = \theta_H\})$. While each individual observes the own taste realization, p is an unknown parameter between 0 and 1. It is assumed that p is the only source of aggregate uncertainty; that is, there exists a bijection between possible values of the parameter $p \in [0, 1]$ and the set $\mathcal{D}$ of states of the economy. To be more precise, the following assumptions are imposed.

Assumption 3.5 The assignment of characteristics to individuals proceeds sequentially. First, there is a skill assignment $w_a : W \to I$. Second, there is an assignment of taste parameters to skill parameters $\theta_a : W \to \Theta$ with image denoted by $\{\theta_w\}_{w \in W}$. The interpretation is that for an individual with productivity level w the taste parameter is given by θ_w. Skill and taste assignments are assumed to satisfy the following properties.

i) For any $p \in [0, 1]$, the profile $\{\theta_w\}_{w \in W}$ is the realization of an *iid* process of random variables $\{\tilde{\theta}_w\}_{w \in W}$.

ii) A Law of Large Numbers applies: almost all realizations of $\{\tilde{\theta}_w\}_{w \in W}$ are such that, for every subinterval $[w_1, w_2] \subset W$,

$$\frac{1}{F(w_2) - F(w_1)} \int_{w_1}^{w_2} \theta_w \, dF \quad = \quad p\theta_H + (1 - p)\theta_L \ .$$

Remark 3.2 In Assumption 3.5 the random taste assignment operates on the set W of possible skill levels and not directly on the set of individuals

I. As both W and I are continua this does not affect the mathematical structure. The results of Al-Najjar (2004) remain applicable.

This specification has the property that in every state of the economy the empirical marginal cross-section distribution of the skill parameter and the empirical marginal cross-section distribution of the taste parameter are independent. Put differently, in every state of the economy, the average taste level is the same on every subinterval of W. A setup that does not use this assumption is the *Two-Class Economy* analyzed in Chapter 4.

3.2.3 Individually incentive compatible allocations and the taxation principle

A tax system is interpreted as the outcome of a mechanism design problem under the restriction that allocations have to be anonymous. An *anonymous allocation* consists of two mappings, a provision rule for the public good,

$$Q : \mathcal{D} \to \mathbb{R}_+, D \mapsto Q(D),$$

and a menu of private goods bundles

$$A : \mathcal{D} \times \Gamma \to \mathbb{R}^l, (D, \gamma) \mapsto A(D, \gamma)\ .$$

Remark 3.3 Following Guesnerie (1995) two aspects of anonymity can be distinguished. There is *recipient anonymity* as the private goods bundle dedicated to an individual depends only on that individual's characteristics but not on the index i. In addition, there is *anonymity in influence*. Neither the menu $\{A(D,\gamma)\}_{\gamma \in \Gamma}$ nor the provision level $Q(D)$ change in response to a permutation of $\{\gamma_i\}_{i \in I}$ that leaves the cross section distribution D unaffected.[6]

[6] Guesnerie (1995) argues that the consideration of anonymous allocations contains no

Definition 3.1 An anonymous allocation is said to be *individually incentive compatible (I-IC)* if

$$\forall D, \forall \hat{\gamma}, \forall \gamma : \quad U(Q(D), A(D, \gamma), \gamma) \geq U(Q(D), A(D, \hat{\gamma}), \gamma) \ .$$

It is *feasible* if for all $D \in \mathcal{D}$, the collection $[Q(D), \{A(D, \gamma)\}_{\gamma \in \Gamma}]$ belongs to the set of feasible allocations Z.

The individual incentive compatibility conditions are stated for a given D; that is, they restrict the possibility for a differential treatment of individuals only within a given cross-section distribution of characteristics. They do not place constraints on the ability of an allocation to specify different outcomes for different members of $\mathcal{D}$. This is due to the fact that in a continuum economy any one individual has a mass of zero and hence does not affect the distribution of characteristics; i.e. there is no impact on the state of the world as perceived by the mechanism designer. In combination with the postulate of anonymity this implies in particular, that no single individual has an impact on public good provision.

Remark 3.4 It is possible to prove a revelation principle for anonymous allocations. Accordingly, the set of anonymous allocations which are implementable as the outcome of some *anonymous game*[7] in which each individual has a dominant strategy coincides with the set of *I-IC* allocations. A more

loss of generality if the profile of characteristics $\{\gamma_i\}_{i \in I}$ is viewed as the realization of an *iid* process of random variables. In that case there is no correlation among individual characteristics that a mechanism designer could potentially exploit.

[7] An anonymous game is defined by the property that a player's payoff depends on the own action and the own characteristics, while the actions chosen by other players only enter via their empirical distribution. More details can be found in Kalai (2004).

precise statement and a proof of this revelation principle can be found in the Appendix.

If one seeks for allocations that can be reached via some anonymous game form, then Remark 3.4 allows to restrict attention to anonymous allocations which are *I-IC* and feasible. Moreover, as the following proposition claims those allocations have the property of being *decentralizable*.

Definition 3.2 An anonymous allocation $[Q, A]$ is called *decentralizable* if there exists a collection of budget sets $\{B(D)\}_{D \in \mathcal{D}}$ such that

$$\forall \gamma, \forall D : \quad A(D, \gamma) \in \quad \text{argmax}_{X \in B(D)} \quad U(Q(D), X, \gamma)$$

Proposition 3.1 (Taxation Principle) An anonymous allocation is *I-IC* and feasible if and only if it is a decentralizable.

A proof can be found in Hammond (1979). According to the *taxation principle*, any *I-IC* and feasible allocation has the property of being decentralizable via a budget set $B(D)$ that is common for all individuals in the economy, and vice versa. Consequently, the set of decentralizable allocations is the relevant object for a study of tax systems. Any tax system generates a decentralized allocation, where the budget set $B(D)$ is shaped by the available tax instruments. The final allocation then results from the solution of the utility maximization problems that individuals face under the given tax system. In reverse direction, the *taxation principle* implies that to each *I-IC* allocation one can find a corresponding tax system – implicitly defined as the set of tax instruments that generate the set $B(D)$.
In the theory of optimal income taxation the *taxation principle* takes a more concise form as illustrated by the following example.

Example 3.2 Suppose that $A = (C, Y)$ as in Example 3.1. An anonymous allocation $[Q, C, Y]$ is said to be *feasible* if

$$\forall D: \quad \int_\Gamma Y(D,\gamma) - C(D,\gamma)\, dD \quad = \quad K(Q(D)) \; ,$$

where $K(\cdot)$ is a cost function that captures the resource requirement of public good provision. The allocation $[Q, C, Y]$ is said to be *decentralizable by an income tax*, if there exists a function $T : \mathcal{D} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ such that for all D and for all γ: Consumption equals after tax income: $C(D,\gamma) = Y(D,\gamma) - T(D, Y(D,\gamma))$, individuals choose a utility maximizing level of income subject to the given income tax schedule,

$$Y(D,\gamma) \quad \in \quad \text{argmax}_Y \quad U(Q(D), Y - T(D,Y), Y, \gamma) \; ,$$

and the public sector budget constraint is satisfied,

$$\int_\Gamma T(D, Y(D,\gamma)) dD \quad = \quad K(Q(D)) \; .$$

For this environment the *taxation principle* takes the following form: An anonymous allocation is *I-IC* and feasible if and only if it is *decentralizable by an income tax*. A proof can be found in the Appendix.

3.2.4 Why individual incentive compatibility is not enough under aggregate uncertainty

As has been emphasized in the previous subsection, the *I-IC* constraints are stated for a given cross-section distribution of characteristics D. They address the *screening problem*, dealing with the question to what extent a differential treatment of individuals with different characteristics is possible if information on these characteristics is private. However, next to the *screening problem* an allocation has to solve a problem of *information aggregation* as the actual distribution D is not commonly known but has to be deduced from the profile of reports $\{\hat{\gamma}_i\}_{i\in I}$ in the revelation game.

One could take the view that this problem of *information aggregation* is solved trivially as a by product of *I-IC* because the economy under consideration is large. No individual has an impact on the mechanism designer's perception of D, and hence there is no reason to hide individual characteristics provided that *I-IC* is ensured.[8] Put differently, as no individual has an impact on public good provision and the shape of the tax system $B(D)$, individuals cannot do better than undertaking a utility maximizing choice taking $B(D)$ as given.

This paper, however, takes a different view. The requirement of *I-IC* still leaves room for collective manipulations which exploit the fact that a subset of agents with positive mass can affect the perceived distribution of characteristics. This is most easily demonstrated if the utility function U is additively separable. In this case *I-IC* cannot ensure that individuals reveal their taste parameter truthfully.

Lemma 3.1 Suppose Assumption 3.1 holds. An anonymous allocation is *I-IC* if and only if it satisfies the following properties:

i) The *no discrimination of taste (NDT)* property:

$$\forall D, \forall w, \forall \theta, \forall \theta' : \quad u(A(D,\theta,w),w) = u(A(D,\theta',w),w) \ .$$

ii) The *individual revelation of productivity (I-RP)* property:

$$\forall D, \forall \theta, \forall w, \forall w' : \quad u(A(D,\theta,w),w) \geq u(A(D,\theta,w'),w) \ .$$

[8] Even ex post, after the actual D has become known, no individual would want to revise his announcement, if hypothetically given the opportunity to do so. This relationship between dominant strategy implementation and ex post incentive compatibility is discussed further in Bergemann and Morris (2005) and Chung and Ely (2003, 2004).

The Lemma follows from the fact that individuals take D and hence the level of public good provision as given. Due to the separability assumption, the utility contribution of the public good vanishes from individual incentive compatibility constraints. In particular, this implies that *I-IC* conditions become independent of taste parameters. Consequently, a tax system uses individual differences in productivity as a screening device and leaves all individuals indifferent regarding possible taste announcements.

The application discussed below shows that this creates an opportunity for collective manipulations of taste announcements that induce a state perception $\hat{D} \neq D$ under which the deviating group of individuals achieves a preferred treatment.

3.2.5 The Example continued

Reconsider the economy described in subsection 3.2.2. Assume that the utility function satisfies Assumption 3.1 and, moreover, takes the following quasi-linear form

$$U \;=\; \theta Q - \frac{t}{w} \,,$$

where t is the individual's contribution to the cost of public good provision. This utility specification captures the idea that less able individuals suffer from a larger utility loss if forced to generate the income that is needed to meet a given payment obligation t.

While U is the cardinal utility function that is relevant for welfare assessments its ordinal properties are equivalently represented by the following monotone transformation $V := wU = \theta w Q - t$. I refer to to the term θw as the *effective valuation* of the public good by an individual with characteristics (θ, w).

Recall the information structure specified in subsection 3.2.2. Each state of the economy corresponds to a realization of the parameter p that determines the average valuation of the public good $p\theta_H + (1-p)\theta_L$. An anonymous allocation is hence represented by a provision rule for the public good $Q : p \mapsto Q(p)$ and a payment scheme $t : (p, \theta, w) \mapsto t(p, \theta, w)$ that specifies for each state of the economy the contribution of an individual with characteristics (θ, w) to the cost of public good provision.

A straightforward application of Lemma 3.1 yields the observation that an allocation $[Q, t]$ is *I-IC* if and only if, for each p all individuals have the same payment obligation; i.e. for all p and for all (θ, w) and all (θ', w'), $t(p, \theta, w) = t(p, \theta', w')$. Moreover, assuming that the cost of public good provision is given by a strictly increasing and strictly convex cost function $K : Q \mapsto K(Q)$ and adding a resource constraint yields the result that $[Q, t]$ is *I-IC* and *feasible* if and only if the payment scheme t prescribes *equal cost sharing*; that is, for all p and for all (θ, w), $t(p, \theta, w) = K(Q(p))$.[9]

These observations allow to represent an individual's assessment of an allocation rule $[Q, t]$, which is budgetary feasible and incentive compatible, in the following reduced form that depends only on the provision rule Q,

$$V(p, \theta, w) \;:=\; \theta w Q(p) - K(Q(p)) \, . \tag{3.1}$$

In what follows, I consider the choice of an optimal *I-IC* and *feasible* allocation by a benevolent utilitarian planner. The planner evaluates an allocation from the ex ante perspective, i.e. before the actual value of p is known. For simplicity, I assume that the planner takes the actual state of the economy p to be the realization of a random variable $\tilde{p}$ that is uniformly distributed on the unit interval $[0, 1]$.[10] Using the Law of Large Numbers in Assumption 3.5,

[9] A more detailed derivation can be found in Chapter 2.

[10] Throughout I do not impose a common prior assumption. Only the prior beliefs of the mechanism designer are specified.

this implies that expected utilitarian welfare from the ex ante perspective is given by

$$
\begin{aligned}
EW &:= \int_0^1 \left\{ \Big[p\theta_H + (1-p)\theta_L\Big] Q(p) - \left[\int_{\underline{w}}^{\bar{w}} \frac{f(w)}{w}\, dw\right] K(Q(p)) \right\} dp \\
&= \lambda \int_0^1 \{\bar{v}(p)Q(p) - K(Q(p))\}\, dp \;,
\end{aligned}
$$

where $\lambda := \int (1/w) f(w) dw$ is an index of the marginal welfare effect of the cost of public good provision under equal cost sharing and

$$\bar{v}(p) \;:=\; \frac{p\theta_H + (1-p)\theta_L}{\lambda}$$

is the *effective utilitarian valuation* of the public good.

I will now show that if EW is maximized under the requirements of *I-IC* and feasibility only, then the resulting allocation is vulnerable to manipulative collective actions by groups of individuals who oppose the decision on public good provision. To see this consider the provision rule $Q^* : p \mapsto Q^*(p)$ that is chosen by a utilitarian planner who maximizes EW pointwise; i.e. who maximizes the expression $\bar{v}(p)Q(p) - K(Q(p))$ for every $p \in [0,1]$. This provision rule is characterized by a continuum of first order conditions

$$\forall p : \;\; \bar{v}(p) \;=\; K'(Q^*(p)) \;.$$

Under Q^* individual preferences over the *"announced state of the world"* can be represented by the following indirect utility function,

$$V^*(p,\theta,w) \;:=\; \theta w Q^*(p) - K(Q^*(p)) \;.$$

It is easily verified that

$$V_p^*(p,\theta,w) \;=\; Q^{*\prime}(p)\Big(\theta w - \bar{v}(p)\Big) \;\; \begin{cases} < 0 & \text{if} \;\; \theta w < \bar{v}(p) \;, \\ = 0 & \text{if} \;\; \theta w = \bar{v}(p) \;, \\ > 0 & \text{if} \;\; \theta w > \bar{v}(p) \;. \end{cases}$$

That is, under provision rule Q^* an individual prefers a larger level of p – or equivalently a larger level of public good provision – if and only if the own *effective valuation* exceeds the *effective utilitarian valuation.* Likewise an individual with an *effective valuation* below the average prefers to have a lower quantity of the public good.

These observations imply that groups of individuals would refuse to reveal their true taste parameters if they could thereby affect the mechanism designer's perception of p. To see this, consider the set of individuals with a low taste realization and a high skill level who have an effective valuation close to $\theta_L \bar{w}$. Moreover, for the sake of concreteness, assume that these individuals share the belief that p is very low.[11] Put differently, these individuals believe that a vast majority has a low taste realization and that, as a consequence, their own effective valuations lie above the effective utilitarian valuation. Hence, under Q^*, this group of individuals expects that the quantity of the public good is too low and would be happy if the mechanism designer had a larger perception of p. But this implies that these individuals are better off if they collectively announce a high taste realization and thereby manipulate the perceived state of the economy.

These considerations highlight the following issues: *First*, a collective deviation from the truth may be beneficial for a subset of agents. *Second* such a collective deviation is not prevented by individual incentive compatibility. Given that all high skilled individuals lie about their taste parameter, there is no incentive for an isolated high skilled individual to reveal his taste parameter truthfully. Due to the *NDT* property, this is a systematic feature.

[11]This means that ex interim these individuals have prior beliefs that put a lot of probability mass on values of p which are close to zero.

With separable preferences a collective deviation involving taste parameters is not undermined by individual incentives.[12]

3.3 Collective Incentive Compatibility

As the discussion in the preceding subsection has shown, the requirement of *I-IC* is not sufficient to ensure that an allocation is able to fulfill the task of information aggregation. Under *I-IC* incentives for a collective manipulation of the mechanism designer's perception of the distribution D are not eliminated. In the following the notion of a *collectively incentive compatible (C-IC)* tax system is introduced that does not suffer from this problem.

More specifically, the definition of *collective incentive compatibility* that is given below requires that truth-telling is a dominant strategy for each conceivable coalition. The main advantage of this approach is that the analysis of coalition formation does not require assumptions on the prior beliefs of individuals. Moreover, as will be explained below, the focus on dominant strategies implies that coalition formation can be analyzed as if individuals had complete information on the state of the economy.

Before the definition of a *C-IC tax system* can be stated, I need to define a *coalition* and a *subcoalition*. For reasons that will become clear, I require that any potentially manipulating subset of agents must have a fixed minimal size $\epsilon > 0$, where ϵ can be arbitrary small. Moreover, a subcoalition J' is a subset of a given coalition J that excludes at least an ϵ - mass of individuals from J.

[12] A further example for the vulnerability of an optimal *I-IC* and *feasible* allocation is found in Chapter 4.

Definition 3.3 A *coalition* J is a subset of agents with $\mu(J) \geq \epsilon$, for some fixed but arbitrary $\epsilon > 0$. A *subcoalition* J' of coalition J is a coalition with the properties $J' \subset J$ and $\mu(J') \leq \mu(J) - \epsilon$.

Two implications of this definition, that are used below, are the following:

i) A coalition J with $\epsilon \leq \mu(J) < 2\epsilon$ possesses no subcoalition.

ii) Consider a chain $\ldots \subset J''' \subset J'' \subset J' \subset J$ resulting from a successive formation of subcoalitions. Any such chain has a finite length.

The following notation is needed to describe the potential impact of a coalition on the perceived distribution of characteristics. Denote by $\mathcal{J}_\epsilon$, with typical element J, the set of subsets of I which satisfy $\mu(J) \geq \epsilon$. Denote the true profile of characteristics in J by $\gamma_J := \{\gamma_j\}_{j \in J}$. Denote the reported profile by $\hat{\gamma}_J := \{\hat{\gamma}_j\}_{j \in J}$. Let the actual distribution of characteristics in the economy be $D \in \mathcal{D}$. Denote the cross section distribution of announcements induced by $\hat{\gamma}_J$ if all individuals not in J report truthfully by $\hat{D}(\hat{\gamma}_J, D) \in \Delta_\Gamma$, where Δ_Γ is the set of cdfs with domain Γ.[13]
Consider a coalition J with $\mu(J) \geq 2\epsilon$. Suppose that J induces state perception $\hat{D}(\hat{\gamma}_J, D)$ via the profile of announcements $\hat{\gamma}_J$. Suppose that the members of a subcoalition J' of J deviate from this profile and report instead according to $\tilde{\gamma}_{J'} \neq \hat{\gamma}_{J'}$. The induced announced distribution of characteristics is denoted by $\hat{D}(\tilde{\gamma}_{J'}, \hat{\gamma}_{J \setminus J'}, D)$.

Definition 3.4 A coalition J is said to *manipulate* an alloation if there exists $D \in \mathcal{D}$, and $\hat{\gamma}_J \neq \gamma_J$ with the following properties:

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

[13] Assumptions 3.3 and 3.4 imply that $\mathcal{D} \subset \Delta_\Gamma$ and $\mathcal{D} \neq \Delta_\Gamma$.

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

$$U(Q(\hat{D}), A(\hat{D}, \hat{\gamma}_j), \gamma_j) \quad > \quad U(Q(D), A(D, \gamma_j), \gamma_j) \; .$$

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

$$U(Q(\hat{D}), A(\hat{D}, \hat{\gamma}_j), \gamma_j) \quad = \quad U(Q(\hat{D}), A(\hat{D}, \gamma_j), \gamma_j) \; .$$

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

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

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

A peculiarity of Definition 3.4 is that collective stability of a coalition J is defined with reference to the collective stability of a subcoalition $J' \subset J$. The requirement of a minimal size for coalitions and subcoalitions ensures that these notions can be traced back to the collective stability of a set of "smallest" coalitions, those with mass between ϵ and 2ϵ.[14]

The requirement of collective incentive compatibility ensures that the allocation $[Q, A]$ can be implemented as the outcome of an anonymous revelation game in such a way that for each coalition J truth-telling is a dominant strategy in the following sense: for any profile of announcements of individuals not in J, truth-telling is the best stable collective announcement for individuals in J.[15] In Remark 3.4 it has been claimed that *I-IC* of an allocation is equivalent to the possibility to implement it as the outcome of an anonymous revelation game in which each individual possesses a dominant strategy. The requirement of *C-IC* is hence commensurate to *I-IC* in the sense that both ensure implementability in dominant strategies.

The interpretation of these requirements in terms of admissible tax systems is the following. According to the *taxation principle* in Proposition 3.1 the

[14]Bernheim et al. (1986) introduce the notion of a *coalition-proof* Nash-equilibrium for games with a finite number of players. They provide a recursive definition based on a definition of coalition-proofness for games with only one player. The above definitions of *stability* are an adaption of this idea for the present setup.

[15]Alternatively, *C-IC* can be framed as a *robustness*-requirement that ensures incentive compatibility of collective actions irrespective of the prior beliefs of individuals in the economy; see Bergemann and Morris (2005) or Kalai (2004).

requirements of anonymity and *I-IC* are equivalent to the existence of a tax system that can be used to decentralize an allocation. However, decentralization via the budget set $B(D)$ presumes that the actual state D has already been determined. The additional requirement of *C-IC* ensures that this information is indeed available; that is, under *C-IC* the tax system does not rely on information that creates a scope for collective manipulations by groups of individuals. Put differently, allocations that are *I-IC* and *C-IC* imply the existence of a tax system and simultaneously allow for information aggregation.

As a final comment, one might take the view that the requirement of *C-IC* is too strong in the sense that there exist alternative ways of achieving a non-manipulable allocation. For instance, a mechanism designer could use use *"off-the-equilibrium rewards"* for subcoalitions to destabilize potential coalitions. To illustrate this, suppose that in state $D \in \mathcal{D}$, coalition J would want to induce state perception $\hat{D} \in \mathcal{D}$ using the false announcements in $\hat{\gamma}_J$. Now suppose that the mechanism designer rewards a further deviation of a subcoalition $J' \subset J$ to some announced distribution $\tilde{D}$, where $\tilde{D}$ does not belong to the set of feasible states $\mathcal{D}$. Thereby the initial manipulation of coalition J is undermined.[16] Moreover it is undermined in a way that is not costly in terms of the welfare properties of the final allocation because the outcome promised to individuals in J' under $\tilde{D}$ is not part of an equilibrium allocation. Hence, an implicit assumption underlying the requirement of *C-IC* is that such *"off-the-equilibrium tax systems"* that only serve to destroy collective manipulations of *"equilibrium tax systems"* can not be used. While this entails a loss of generality, it still seems to be a reasonable way of modeling tax systems.

[16] A similar reasoning can be found in Boylan (1998).

3.3.1 Related Literature

The requirement of *C-IC* uses the notion of a *coalition-proof Nash equilibrium* that has been developed by Bernheim et al. (1986). These authors propose a refinement of the Nash equilibrium concept for games of complete information. As in this paper, the incentives coalition members face individually must not conflict with the action profile used by the coalition and moreover a conceivable collective manipulation has to be such that it does not provoke the formation of a further subcoalitions that depart from the initial coalitional plan, where a potentially deviating subcoalition again has to meet these stability requirements.

To relate their solution concept for games of complete information to the setting of this paper, the requirement of *C-IC* can be interpreted as follows. Suppose that, for some reason, the actual state of the economy D is commonly known among all individuals and that the mechanism designer is the only uninformed party. Still, the mechanism designer uses the revelation game to learn the actual state of the economy and to choose the level of public good provision $Q(D)$ and the menu of private goods bundles $B(D)$. The revelation game has thus become a game of complete information.[17] Moreover, each $D \in \mathcal{D}$ gives rise to a different complete information game. With this interpretation the requirement of *C-IC* can be stated as follows. *C-IC* holds if and only if in each complete information game truth-telling is a stable best response for each coalition, given that all individuals outside the coalition tell the truth.

The insistence on stability of coalitions with respect to the formation of subcoalitions distinguishes the present paper from some recent contributions to

[17] Moore (1992) provides a survey of implementation problems in environments with complete information.

the literature on mechanism design problems under the possibility of coalition formation. In a series of papers Laffont and Martimort (1997, 1999, 2000) incorporate a sequential Bayesian game of coalition formation into a mechanism design problem. These authors however only consider collective manipulations by the grand coalition. By contrast, Demange and Guesnerie (2001) allow for the formation of coalitions smaller than the grand coalition. As Laffont and Martimiort they do not require stability with respect to the formation of subcoalitions. Instead they are concerned with concepts of the core in games of incomplete information without aggregate uncertainty.

3.3.2 Aggregate Uncertainty and Undetectability

Recall how aggregate uncertainty has been formalized by Assumptions 3.3 and 3.4. For any $w \in W$, the "share" of individuals with productivity parameter w is commonly known. By contrast, the "share" of individuals with a taste parameter θ among those with productivity w, is not commonly known for all $w \in W$. Finally, those properties of the joint distribution of taste and skill parameters that are commonly known, determine the structure of the set $\mathcal{D}$.

The *undetectability* requirement in the above definition of a *C-IC* tax system precludes the formation of coalitions which induce an announced distribution of characteristics that does not belong to $\mathcal{D}$. Implicitly it is thus assumed, that the mechanism designer can effectively deter those collective manipulations for which it becomes obvious that some set of agents must have been deviating from the truth.[18]

[18] Note that even if a manipulation becomes apparent, the manipulating individuals are not yet identified. The above definition hence implicitly relies on the assumption, that the mechanism may punish *all* individuals harshly in response to an obvious collective lie.

The difficulty of achieving *C-IC* depends to a large extent on the assumptions on the feasible set $\mathcal{D}$ and the mechanism designer's ability to detect collective manipulations. To illustrate this, the application specified in subsections 3.2.2 and 3.2.5 is discussed once more.

3.3.3 The Example continued

Again consider the example discussed in subsections 3.2.2 and 3.2.5. Recall that it is assumed to be commonly known, that the average taste realization $p\theta_H + (1-p)\theta_L$, is the same on every subset of W. Aggregate uncertainty stems only from the fact that p itself is an unknown parameter.

One can take the view that if the empirical taste and the empirical skill distribution satisfy this property of independence almost surely, then basically any collective manipulation is detectable. Whenever agents from a particular part of the skill distribution form a manipulating coalition – while all other agents stick to the truth – this induces an announced distribution of characteristics which is inconsistent with the commonly known fact that the average taste level is the same on every subset of W. Consequently, an undetectable manipulation has to be such that the average taste level is affected on every subinterval of W in the same way. This basically requires that the whole set of agents I is willing to undertake a collective manipulation. The only coalition which might potentially undermine an allocation is thus the so called *grand coalition* consisting of all agents. This is a perfectly consistent view on *undetectability*. It is formalized below.

In addition, I define an alternative which is such that the mechanism designer cannot impose as much discipline on potential coalitions. As the realization of taste parameters is governed by an *iid* process of random variables, it

---

This in turn implies, that no coalition will consider such a collective plan.

can in principle happen that the average taste level is different for different subintervals of W. Even though such an event has probability zero it is not excluded from the support of the stochastic process $\{\tilde{\theta}_w\}_{w \in W}$. This distinction between supported outcomes and those which arise with strictly positive probability allows for two different versions of the *undetectability* requirement:

Definition 3.5 Consider the application specified in subsections 3.2.2 and 3.2.5.

i) A collective manipulation is *weakly undetectable* if the induced skill distribution is given by F.

ii) A collective manipulation is *strictly undetectable* if the induced skill distribution is given by F and the induced taste distribution is such that average taste level is the same on every subinterval of the skill distribution.

Below, in section 3.5, the set of allocations which are not only *I-IC* but also *C-IC* is characterized for this environment. As will become clear, which version of *undetectability* is used, has a huge impact on the set of admissible allocations.

3.4 Separability

In this section the set of allocations that are feasible, *I-IC* and *C-IC* is analyzed under the assumption that the utility function U is additively separable. This allows to establish a property which proves very useful in applications: The different incentive concerns can be separated. The requirement of *I-IC* deals with the resolution of pure assignment uncertainty in the profile of skill parameters; i.e. it allows to solve the screening problem of identifying individual skill levels within a given cross-section distribution F. The postulate of *C-IC* is concerned with problem of information aggregation which arises due to the aggregate uncertainty in the joint distribution of skill and taste parameters. It turns out that, in order to ensure *C-IC*, it suffices to eliminate incentives for a collective manipulation of taste parameters. There is no need to worry about collective manipulations of reported skill parameters.

Definition 3.6 A *utility allocation* is a mapping $\tilde{U} : (D, \gamma) \mapsto \tilde{U}(D, \gamma)$. A utility allocation $\tilde{U}$ is said to be *implementable* if there exists an anonymous allocation $[Q, A]$, which is feasible, *I-IC* and *C-IC* and such that:

$$\forall D, \forall \gamma : \quad \tilde{U}(D, \gamma) \;=\; U(Q(D), A(D, \gamma), \gamma) \; .$$

It will prove helpful to have an own terminology for coalitional manipulations which are based on a false report of taste parameters but which are truthful with respect to the reported skill parameters. A typical message profile of a manipulating coalition J which is such that, $\forall j \in J$, the reported skill parameter $\hat{w}_j$ is equal to the true skill parameter w_j is henceforth called a *partial taste manipulation* and denoted by $\hat{\gamma}^p_J$. To emphasize that some *manipulation* $\hat{\gamma}_J$ is not *partial*, I write $\hat{\gamma}_J = [\hat{w}_J, \hat{\theta}_J]$ with $\hat{w}_J \neq w_J$.

Definition 3.7

i) A coalition J is said to possess a *partial taste manipulation* if there exists $D \in \mathcal{D}$ and an *undetectable* partial manipulation $\hat{\gamma}^p_J$ that induces a state perception $\hat{D}(\hat{\gamma}^p_J, D)$, which makes all members of J strictly better off relative to D (*unanimity*), prescribes for all its members taste announcements which are individually a best response in conjunction with a truthful skill announcement under state perception $\hat{D}(\hat{\gamma}^p_J, D)$ (*individual stability*), and is not threatened by further partial taste manipulations of subcoalitions, which satisfy all these requirements (*collective stability*).

An allocation is said to have the *collective revelation of taste (C-RT)* property if there does not a exist a coalition with a partial taste manipulation.

ii) A utility allocation is *partially implementable* if there exists an anonymous allocation $[Q, A]$, which is feasible, *I-IC*, has the *C-RT* property and is such that:

$$\forall D, \forall \gamma : \quad \tilde{U}(D, \gamma) \quad = \quad U(Q(D), A(D, \gamma), \gamma) \ .$$

Obviously, if an allocation is *C-IC*, then it has also the *C-RT* property. As a consequence, the set of implementable utility allocations is a subset of the set of partially implementable allocations. The following Lemma shows that the converse inclusion holds true as well. Hence, it justifies an analysis of allocations which possess only the *C-RT* property.

Lemma 3.2 Under assumptions 3.1 - 3.4, the set of implementable utility allocations is equal to the set of partially implementable utility allocations.

The proof is based on the observation that under aggregate stability with respect to the distribution of skill parameters, any conceivable undetectable collective manipulation which involves both taste and skill parameters can be mimicked by a partial manipulation which involves only reported taste parameters. Intuitively, any manipulation has to be undetectable. Hence, whenever a subset of agents J manipulates via some $\hat{\gamma}_J = [\hat{w}_J, \hat{\theta}_J]$ with $\hat{w}_J \neq w_J$, this manipulation has to be such that the resulting distribution of announcements $\hat{D}$ has a marginal skill distribution which is equal to F. But this implies that coalition J can induce $\hat{D}$ as well by a suitably chosen partial taste manipulation. As a consequence, it suffices to exclude the possibility of partial taste manipulations in order to establish the *C-IC* property.

The results in Lemmas 3.1 and 3.2 imply that, under assumptions 3.1-3.4, attention can be restricted to the set of feasible allocations which satisfy *I-RP*, *NDT* and *C-RT*. These observations are summarized in the following Theorem.

Theorem 3.1 Under assumptions 3.1-3.4, any implementable utility allocation is also implementable via an allocation $[Q, A]$ that satisfies the following properties: *I-RP*, *NDT*, *C-RT* and *feasibility*.

The *C-RT* constraints may seem rather opaque at the present level of abstraction. It is not obvious how to represent them by a well-defined set of constraints that could, for instance, be included in an exercise of solving for an optimal constrained efficient allocation. To illustrate the impact of the *C-RT* property the next section returns once more to the application.

3.5 The Example continued

This section returns to the application already discussed in subsections 3.2.2, 3.2.5 and 3.3.3. For this environment the set of allocations that are *feasible*, *I-IC* and *C-IC* is explicitly characterized in the following. This finally allows to solve for the *optimal* allocation that meets all these criteria.

By Theorem 3.1 attention is restricted to allocations $[Q, t]$ that satisfy *I-RP*, *NDT*, *feasibility* and the *C-RT* property. As has already been observed in subsection 3.2.5, the first three requirements are equivalent to the payment scheme t being such that the cost of public good provision are shared equally among all individuals, for every state p of the economy. Consequently, the remaining task is to characterize the provision rules $Q : p \mapsto Q(p)$ that yield the *C-RT* property under such a payment scheme. To achieve this, the two versions of the *undetectability* requirement, that have been introduced in definition 3.5 have to be distinguished.

3.5.1 Undetectability in the strict sense

Recall that under *undetectability* in the strict sense the mechanism designer infers the actual value of the parameter p from a profile of reports $\{\hat{\gamma}_i\}_{i\in I} = \{(\hat{\theta}_i, \hat{w}_i)\}_{i\in I}$ and is able to deter any manipulation such that the reported average taste level is different on different subintervals of W.

Proposition 3.2 Suppose a manipulation is called undetectable if it is *strictly undetectable* in the sense of definition 3.5. Any pair $[Q, t]$ that satisfies *equal cost sharing* also satisfies the *C-RT* property if there do not exist p and p' such that,

$$\forall w, \forall \theta : \quad V(p, \theta, w) \quad > \quad V(p', \theta, w) \; . \tag{3.2}$$

Under *undetectability* in the strict sense basically any provision rule $Q : p \mapsto Q(p)$ is implementable if accompanied by equal cost sharing. The only additional restriction imposed by *C-RT* is that there must not exist a state of the world p such that all individuals unanimously agree that there exists a preferred outcome of the revelation game.[19] This implies in particular, that provision rule $Q^*(p)$ which maximizes EW pointwise, is implementable.

3.5.2 Undetectability in the weak sense

Under *undetectability* in the weak sense, an inconsistency of average taste levels on different subintervals of W is a possible event. The following assumption on the mechanism designer's perception of p does allow for such announcements in the revelation game.

Assumption 3.6 The mechanism designer's perception of p is given by $p = \mu(\{i \mid \hat{\theta}_i = \theta_H\})$, where $\hat{\theta}_i$ is the taste announcement of individual i in the revelation game.

Remark 3.5 Measurability of the set $\{i \mid \hat{\theta}_i = \theta_H\}$ is again ensured with reference to Al-Najjar (2004). In his model of a large economy I is a countable set of infinitely many individuals and the set $\{i \mid \hat{\theta}_i = \theta_H\}$ is measurable with respect to an appropriate generalization of the counting measure.

Assumption 3.6 implies that the mechanism designer chooses the same provision level $Q(p)$ whenever he observes that $\mu(\{i \mid \hat{\theta}_i = \theta_H\})$ is equal to p.[20] Consequently, under assumption 3.6 there is an obvious channel along which

[19] This is a sufficient condition. (3.2) implies that there does not exist a partial taste manipulation for the *grand coalition* of all agents which satisfies *Undetectability*, *Unanimity* and *Individual Stability*.

[20] The provision rule Q can thus be viewed as resulting from a voting procedure. To see this, interpret a high (low) taste announcement as a vote in favor of a large (small) level of public good provision. In this sense, the provision rule $p \mapsto Q(p)$ specifies a provision level for each conceivable vote distribution; see Chapter 2 for more details.

a coalition might manipulate an allocation. Any partial taste manipulation that affects the share of high taste announcements has an effect on the level of public good provision. In particular, this implies that any partial taste manipulation that affects the average taste level on some subinterval of W becomes effective. The following proposition derives the implications of this property for the set of implementable allocations.

Proposition 3.3 Suppose that assumption 3.6 applies. Consider an allocation $[Q, t]$ with equal cost sharing. $[Q, t]$ has the *C-RT* property for any minimal coalition size ϵ if and only if the following properties are satisfied.

i) Q is a non-decreasing function of p.

ii) $V(p, \theta_L, \bar{w})$ is non-increasing and $V(p, \theta_H, \bar{w})$ is non-decreasing in p.

The *"if-part"* in the proposition follows from the observations that, under a non-decreasing provision rule, $V(p, \theta_L, \bar{w})$ is non-increasing in p only if an individual with effective valuation $\theta_L \bar{w}$ always desires a small provision level over a large provision level. This implies that the same is true for any individual with an effective valuation $\theta_L w \leq \theta_L \bar{w}$. As a consequence, no individual with a low taste realization is willing to join a manipulating coalition that attempts to achieve a larger perception of the average taste parameter p, or, equivalently, a larger quantity of the public good. Analogously one shows that no individual with a high taste realization wants to achieve a smaller perception of p. Consequently, even with the opportunity to undertake manipulative collective actions, individuals cannot to better than to reveal their taste parameter.

The proof of the *"only if-part"* is based on the observation that whenever property i) or property ii) is violated, then there exists some small ϵ such that the *C-RT* property fails. For instance, as shown in the appendix, if there exist p' and p with $p' - p = \epsilon$ and $Q(p') < Q(p)$, then there exists a small coalition of individuals with a low taste realization that tries to induce the outcome $Q(p')$ if the true state is p, or a coalition of high taste individuals that aims at $Q(p)$ if the true state is p'. Hence, the *C-RT* property holds for any small ϵ only if properties i) and ii) are fulfilled.

Using Proposition 3.3 it is easily verified that Q^*, the welfare maximizing provision rule under equal cost sharing, is not part of an implementable allocation. As has already been discussed in subsection 3.2.5, there exists a range of small values of p such that the indirect utility function under Q^*, $V^*(p, \theta_L, \bar{w})$, is strictly increasing in p. Analogously one can show that there exists a range of large values of p such that $V^*(p, \theta_H, \underline{w})$ is strictly decreasing in p.

The problem of finding the optimal allocation $[Q, t]$ which maximizes expected welfare and satisfies *equal cost sharing* and, in addition, the *C-RT* property is extensively discussed in Chapter 2.[21] The main result is that an optimal provision rule is characterized by *pooling*; that is, there are several ranges over which an optimal provision rule is constant.

To illustrate what such an optimal deviation from Q^* in the presence of *C-RT* constraints can look like, suppose that the parameters of the model satisfy $\theta_L \bar{w} < \theta_H \underline{w}$. This implies that

$$Q^*(0) \quad < \quad \bar{Q}_L \quad < \quad \underline{Q}_H \quad < \quad Q^*(1) \; ,$$

where $\bar{Q}_L$ is the most preferred provision level of an individual with effective valuation $\theta_L \bar{w}$ under equal cost sharing, $\{\bar{Q}_L\} := \text{argmax}_Q \; \theta_L \bar{w} Q - K(Q)$.

[21] Even though that paper is concerned voting mechanisms for the purpose of information aggregation it arrives at the same characterization of implementable provision rules as Theorem 3.3.

Likewise, $\underline{Q}_H$ is the most preferred provision level of an individual with effective valuation $\theta_H \underline{w}$. The image of provision rule Q^* contains the intervals $[Q^*(0), \bar{Q}_L]$ and $[\underline{Q}_H,\, Q^*(1)]$. However under properties i) and ii) in Proposition 3.3 there can be at most one provision level below $\bar{Q}_L$.[22] Analogously, there is at most one provision level exceeding $\underline{Q}_H$.

As shown in Chapter 2 the following can be an optimal response to these restrictions:

$$Q(p) \quad := \quad \begin{cases} Q^s & \text{for} \quad 0 \leq p \leq \hat{p} \ , \\ Q^{sm} & \text{for} \quad \hat{p} < p < \hat{p}' \ , \\ Q^*(p) & \text{for} \quad \hat{p}' \leq p \leq \tilde{p}' \ , \\ Q^{lm} & \text{for} \quad \tilde{p}' < p < \tilde{p} \ , \\ Q^l & \text{for} \quad \tilde{p} \leq p \leq 1 \ . \end{cases}$$

This provision rule has four pooling levels Q^s, Q^{sm}, Q^{lm} and Q^l and, moreover, over an intermediate range this provision rule coincides with Q^*. The pooling levels Q^s and Q^{sm} are chosen such that $Q^s < \bar{Q}_L < Q^{sm}$ and an individual with effective valuation $\theta_L \bar{w}$ is indifferent between these two provision levels; that is, $\theta_L \bar{w} Q^s - K(Q^s) = \theta_L \bar{w} Q^{sm} - K(Q^{sm})$.[23] Similarly, an individual with effective valuation $\theta_H \underline{w}$ is indifferent between the pooling levels Q^l and Q^{lm}.

It is shown in Chapter 2 that the shape of the optimal provision rule depends on the parameters of the model. If heterogeneity in productivity parameters is relatively small – this is the case if $\underline{w}$ is close to $\bar{w}$ – then an optimal provi-

[22] To see this, suppose to the contrary, that there are p and p' with $Q(p) < Q(p') < \bar{Q}_L$. Then, because of the fact that the function $\theta_L \bar{w} Q - K(Q)$ is single peaked, an individuals with effective valuation prefers $Q(p')$ over $Q(p)$. However, this contradicts property ii) in Proposition 3.3.

[23] Note that if there where p and $p' > p$ such that $Q(p) = Q^s$ and $Q(p') < Q^{sm}$, then property ii) in Proposition 3.3 would be violated.

sion rule has four pooling levels and is rather close to provision rule Q^* which would be optimal without *C-RT* constraints. If, however, heterogeneity with respect to productivity levels is more pronounced one may even end up with a constant provision that does not use any information on taste realizations. These considerations show that the requirement of *C-RT* may have a drastic impact on the optimal provision rule for a public good.

3.6 Appendix

Statement and Proof of Revelation Principle

An *anonymous mechanism* M is a game form consisting of a message space R, a provision rule for the public good and a menu of consumption-income combinations. To describe these functions denote by Δ_R the set of cumulative distribution functions (cdfs) on R and denote a typical element of Δ_R by ρ. An anonymous mechanism is defined by the mappings:

$$Q^M : \Delta_R \to \mathbb{R}_+,\ \rho \mapsto Q^M(\rho)\ ,$$

$$A^M : \Delta_R \times R \to \mathbb{R}^2_+,\ (\rho, r) \mapsto A^M(\rho, r)\ .$$

A *direct anonymous mechanism* $\bar{M}$ is an anonymous mechanism which satisfies $R = \Gamma$ and is summarized by the functions

$$Q^{\bar{M}} : \Delta_\Gamma \to \mathbb{R}_+,\ D \mapsto Q^{\bar{M}}(D)\ ,$$

$$A^{\bar{M}} : \Delta_\Gamma \times \Gamma \to \mathbb{R}^2_+,\ (D, \gamma) \mapsto A^{\bar{M}}(D, \gamma)\ .$$

Note that the domain of a *direct anonymous mechanism* does not coincide with the one of an anonymous allocation defined in the body of the text. The reason is that an anonymous allocation specifies the level of public good provision Q and the menu of consumption-income pairs A only for cdfs which belong to the feasible set $\mathcal{D}$. By contrast, a *direct anonymous mechanism* specifies an outcome of the game for each distinguishable action profile, that is, for each distribution of announcements in Δ_Γ.

Consider the game induced by an anonymous mechanism M. A *strategy* s for an agent assigns a report to each possible value of individual characteristics. Formally:

$$s : \Gamma \to R : r = s(\gamma) \; .$$

Denote the set of possible strategies by S.

The game induced by anonymous mechanism M has an *equilibrium in dominant strategies* if there exists a mapping s^* such that $\forall \gamma$, $\forall \rho \in \Delta_R$ and $\forall s \in S$:

$$U(Q^M(\rho), A^M(\rho, s^*(\gamma)), \gamma) \geq U(Q^M(\rho), A^M(\rho, s(\gamma)), \gamma) \; .$$

In words: Each type γ has a best response $s^*(\gamma)$, which applies independently of the behavior of others, i.e. which is optimal for all $\rho \in \Delta_R$.

An anonymous mechanism M *implements* an anonymous allocation *in dominant strategies* if the game induced by M has an equilibrium in dominant strategy s^* which satisfies $\forall \gamma$ and $\forall D \in \mathcal{D}$:

$$Q(D) = Q^M(\rho^*(D)) \quad \text{and} \quad A(D, \gamma) = A^M(\rho^*(D), s^*(\gamma)) \; ,$$

where $\rho^*(D)$ is the distribution of reports generated by s^* if the state of the economy is D. Put differently $\rho^*(D)$ is the distribution on R induced by the message profile $\{s^*(\gamma_i)\}_{i \in I}$ if the cdf that corresponds to the profile of characteristics in $\{\gamma_i\}_{i \in I}$ is D.

Consider the game induced by a direct anonymous mechanism $\bar{M}$. *Truth-telling* is a strategy defined by $s(\gamma) = \gamma$ for all $\gamma \in \Gamma$. Truth-telling by all agents is an equilibrium in dominant strategies provided that $\forall\gamma, \forall\hat{\gamma}$ and $\forall D \in \Delta_\Gamma$:

$$U(Q^{\bar{M}}(D), A^{\bar{M}}(D,\gamma),\gamma) \geq U(Q^{\bar{M}}(D), A^{\bar{M}}(D,\hat{\gamma}),\gamma)\ . \tag{3.3}$$

An anonymous allocation is *truthfully implementable* in dominant strategies if there exists a direct anonymous mechanism $\bar{M}$ that implements it such that truth-telling is a dominant strategy; i.e. truthful implementation requires that truth-telling by all agents is an equilibrium in dominant strategies in the game induced by $\bar{M}$, and in addition $\forall\gamma$ and $\forall D \in \mathcal{D}$:

$$Q(D) = Q^{\bar{M}}(D) \quad \text{and} \quad A(D,\gamma) = A^{\bar{M}}(D,\gamma)\ . \tag{3.4}$$

Lemma 3.3 An anonymous allocation rule is *I-IC* if and only if it is truthfully implementable.

Proof The if-part follows from substituting the equations in (3.4) into the inequalities in (3.3), for $D \in \mathcal{D}$. This yields the definition of an *I-IC* allocation. To prove the only if-part, suppose that the pair $[Q, A]$ is an *I-IC* anonymous allocation rule. It has to be shown that there exists a direct anonymous mechanism $[Q^{\bar{M}}, A^{\bar{M}}]$ which implements $[Q, A]$. This direct anonymous mechanism has to be such that the incentive structure is preserved, i.e. such that truth-telling is a dominant strategy. It can for instance be constructed as follows. For all $\gamma \in \Gamma$ and all $D \in \mathcal{D}$ choose $[Q^{\bar{M}}, A^{\bar{M}}]$ such that (3.4) holds. For all $D \in \Delta_\Gamma \backslash \mathcal{D}$ and $\gamma \in \Gamma$, let $Q(D) =$ constant and $A(D,\gamma) =$ constant.

■

Proposition 3.4 (Revelation Principle) An anonymous allocation is implementable if and only if it is truthfully implementable.

Proof The if-part is trivial. Suppose $[Q, A]$ is implementable by some mechanism M. Then there exists a function s^* such that $\forall\gamma, \forall\hat{\gamma}, \forall D \in \mathcal{D}$:

$$U(Q^M(\rho^*(D)), A^M(\rho^*(D), s^*(\gamma)), \gamma) \geq U(Q^M(\rho^*(D)), A^M(\rho^*(D), s^*(\hat{\gamma})), \gamma) \; .$$

(In words: In a dominant strategy equilibrium, the following has to be true. The actions prescribed by the equilibrium strategy s^* are such that no type wants to deviate to an action prescribed for another type, taking the distribution over equilibrium actions as given.) and such that $\forall\gamma, \forall D \in \mathcal{D}$:

$$Q(D) = Q^M(\rho^*(D)) \quad \text{and} \quad A(D, \gamma) = A^M(\rho^*(D), s^*(\gamma)) \; .$$

Combining those statements yields the definition of truthful implementability or equivalently of an *I-IC* anonymous allocation.

■

Proof of Taxation Principle in Example 3.2.

"$\Longleftarrow$": Consider a feasible anonymous allocation. Suppose it is an income tax but not *I-IC*. Then there exist γ, $\hat{\gamma}$ and D such that

$$U(Q(D), A(D, \gamma), \gamma) < U(Q(D), A(D, \hat{\gamma}), \gamma) \; .$$

Using that for all γ, $A(D, \gamma) = [Y(D, \gamma) - T(D, Y(D, \gamma)), Y(D, \gamma)]$, this is equivalent to

$$U(Q(D), Y(D, \gamma) - T(D, Y(D, \gamma)), Y(D, \gamma), \gamma)$$

$$< U(Q(D), Y(D, \hat{\gamma}) - T(D, Y(D, \hat{\gamma})), Y(D, \hat{\gamma}), \gamma) \; .$$

But this contradicts that $\forall D, \forall\gamma$:

$$Y(D, \gamma) \quad \in \quad \arg\max_Y U(Q(D), Y - T(D, Y), Y, \gamma)$$

"$\Longrightarrow$": Consider a feasible and *I-IC* allocation and construct T as follows:

i) For any x such that there is D and γ with $Y(D,\gamma) = x$ define $T(D,x)$ by the equation[24]

$$T(D,x) \quad = \quad Y(D,\gamma) - C(D,\gamma) \ .$$

Obviously, this choice ensures that under T, consumption equals after tax income and that budget balance holds.

ii) For all other levels of Y set $T(D,x) = x$.[25]

Now suppose this function T does not satisfy the property that $\forall D, \forall \gamma$:

$$Y(D,\gamma) \quad \in \quad \arg\max\nolimits_Y U(Q(D), Y - T(D,Y), Y, \gamma)$$

Then there exist $\gamma, \hat{\gamma}$ and D such that

$$U(Q(D), Y(D,\gamma) - T(D, Y(D,\gamma)), Y(D,\gamma), \gamma)$$

$$< U(Q(D), Y(D,\hat{\gamma}) - T(D, Y(D,\hat{\gamma})), Y(D,\hat{\gamma}), \gamma) \ .$$

or using that for all γ, $A(D,\gamma) = [Y(D,\gamma) - T(D, Y(D,\gamma)), Y(D,\gamma)]$,

$$U(Q(D), A(D,\gamma), \gamma) < U[Q(D), A(D,\hat{\gamma}), \gamma] \ .$$

This contradicts *I-IC*.

■

[24]Note that his equation uniquely determines $T(D,x)$. If not, one had, for given D, different consumption levels corresponding to the same income requirement; hence a contradiction to individual incentive compatibility, assuming monotonicity of preferences.

[25]It is implicitly assumed that, for any agent, zero consumption implies a utility level of $-\infty$ and that hence the corresponding Y is never chosen, whenever there is an alternative with positive consumption available.

Proof of Lemma 3.1. To proof the only if-part note that, because preferences satisfy Assumption 3.1, the *NDT-U* property is an implication of *I-IC*. Obviously *I-RP* is also an implication of *I-IC*. To prove the if-part, suppose an allocation rule, such that the *NDT-U* and the *I-RP* property hold, is not *I-IC*. Then there exist (θ, w) and $(\hat{\theta}, \hat{w})$ and D such that $u(A(D, \theta, w), w) < u(A(D, \hat{\theta}, \hat{w}), w)$. Using *NDT-U* and *I-RP* one has:

$$u(A(D, \hat{\theta}, \hat{w}), w) \;\; = \;\; u(A(D, \theta, \hat{w}), w) \;\; \leq \;\; u(A(D, \theta, w), w) \; .$$

Hence, a contradiction.

■

Proof of Lemma 3.2. It has to be shown that any partially implementable utility allocation is implementable. Suppose to the contrary that there exists a partially implementable utility allocation $\tilde{U}$, which is not implementable.

i) Denote by $[Q, A]$ the feasible, *I-IC* and *C-RT* allocation which partially implements $\tilde{U}$. By hypothesis $\tilde{U}$ is not implementable. Hence, there must exist D and a coalition J and a manipulation $\hat{\gamma}_J = [\hat{w}_J, \hat{\theta}_J]$ with $\hat{w}_J \neq w_J$ such that, by *Undetectability*, $\hat{D}(\hat{\gamma}_J, D) \in \mathcal{D}$ and, $\forall i \in J$, by *Individual Stability* and *Unanimity*

$$\begin{aligned} & v(Q(\hat{D}(\hat{\gamma}_J, D)), \theta_i) + u(A(\hat{D}(\hat{\gamma}_J, D), \hat{\theta}_i, \hat{w}_i), w_i) \\ & = v(Q(\hat{D}(\hat{\gamma}_J, D)), \theta_i) + u(A(\hat{D}(\hat{\gamma}_J, D), \theta_i, w_i), w_i) \qquad (3.5) \\ & > v(Q(D), \theta_i) + u(A(D, \theta_i, w_i), w_i) \end{aligned}$$

and such that *collective stability holds.*

ii) *Claim.* The coalition J can induce the announced distribution $\hat{D}(\hat{\gamma}_J, D)$ also via some partial taste manipulation $\hat{\gamma}^p_J$.

Proof. There is aggregate stability with respect to the marginal distribution of skill parameters. Hence, any undetectable manipulation $\hat{\gamma}_J$ with $\hat{w}_J \neq w_J$ has to be consistent with the commonly known skill distribution F. The manipulation of J presumes that all individuals not in J reveal their characteristics truthfully. Hence, to be undetectable, $\hat{\gamma}_J$ has to be such that the distribution of skill announcements within coalition J is equal to the true skill distribution within coalition J. But this implies that the outcome achieved via $\hat{\gamma}_J$ is also induced if all members of J reveal their skill parameter truthfully and choose a suitable profile of announced taste parameters. I.e. for given $\hat{\gamma}_J$ with $\hat{w}_J \neq w_J$, there exists $\hat{\gamma}^p_J$ with $\hat{w}_J = w_J$ such that $\hat{D}(\hat{\gamma}_J, D) = \hat{D}(\hat{\gamma}^p_J, D)$.

iii) *Claim.* The partial taste manipulation $\hat{\gamma}^p_J$ defined with reference to $\hat{\gamma}_J$ in ii) satisfies *Individual Stability* and *Unanimity.*

Proof. $\hat{\gamma}^p_J$ is a partial taste manipulation. Under the separability assumption 3.1, *I-IC* implies the *NDT* property. Hence, any partial manipulation satisfies *Individual Stability.* *Unanimity* follows from $\hat{D}(\hat{\gamma}_J, D) = \hat{D}(\hat{\gamma}^p_J, D)$ and the inequality in (3.5).

iv) If $\hat{\gamma}^p_J$ was collectively stable with respect to partial taste manipulations by subcoalitions of J, then this would contradict, the *C-RT* property of utility allocation $\tilde{U}$. Hence, I assume in the following that $\hat{\gamma}^p_J$ is not collectively stable with respect to partial taste manipulations by subcoalitions of J. I.e. if the true distribution of characteristics in the economy is D and coalition J has induced the announced distribution $\hat{D}(\hat{\gamma}_J, D)$, then there exists a subcoalition J' of J with a partial taste manipulation

$\tilde{\gamma}^p_{J'} \neq \hat{\gamma}^p_{J'}$ which induces a state perception $\hat{D}(\tilde{\gamma}^p_{J'}, \hat{\gamma}^p_{J \setminus J'}, D) \in \mathcal{D}$ (*Undetectability*), which is strictly preferred by all members of J' relative to $\hat{D}(\hat{\gamma}_J, D)$ (*Unanimity*), is *individually stable* and does not provoke partial taste manipulations by further subcoalitions (*collective stability*).

v) *Claim.* It has to be true that the partial taste manipulation $\tilde{\gamma}^p_{J'}$ by subcoalition J' characterized in iv) is not collectively stable with respect to all manipulations $\bar{\gamma}_{J''} \neq \tilde{\gamma}_{J''}$ with $\bar{w}_J \neq w_J$ by subcoalitions J'' of J'.
Proof. Suppose otherwise. Then this partial taste manipulation $\tilde{\gamma}^p_{J'}$ could be used by the set J' to manipulate the initial manipulation of allocation $[Q, A]$ by coalition J via $\hat{\gamma}_J$ in step i), thereby contradicting the *collective stability* of this manipulation.

vi) The reasoning established so far has a recursive structure: The starting point was in step i) an allocation $[Q, A]$, which is not vulnerable by partial taste manipulations but by a joint manipulations $\hat{\gamma}_J$ of both taste and skill parameters. In steps ii)-v) it has been shown that this implies the existence of a subcoalition J' of J which possess a partial taste manipulation $\tilde{\gamma}^p_{J'}$ which does not provoke further partial taste manipulations by subcoalitions of J' but further joint manipulations of both taste and skill parameters by subcoalitions of J'.
Now the reasoning in steps i) - v) can be applied again to show that this implies the existence of a subcoalition J'' of J' which possess a partial taste manipulation but provokes further joint manipulations of both taste and skill parameters by subcoalitions of J'' etc.
However, as a consequence of definition 3.3, any chain of successive formation of subcoalitions has a finite length. Hence, after a finite number of repeated applications of the reasoning in steps i) - v) one ends

up with a situation in which a subcoalition J^x of minimal size – that is, J^x possesses no further subcoalitions – possess a joint manipulations of both taste and skill parameters but not a partial taste manipulation. A last application of steps ii) and iii) then yields a contradiction.

■

Proof of Proposition 3.2. If (3.2) holds, then there is no coalition which is willing to affect the average taste level on all subintervals of W. Any coalition which affects the average taste level only on some subintervals of W is detected.

■

Proof of Proposition 3.3. The proof follows from Lemmas 3.4 – 3.7 below.

Lemma 3.4 Suppose that assumption 3.6 applies. Let the minimal coalition size ϵ be close to zero. If an allocation $[Q, t]$ with equal cost sharing satisfies the *C-RT* property, then $Q(p') \geq Q(p)$ for any pair $p', p \in (0, 1)$ with $\epsilon < p' - p \leq 2\epsilon$.

Proof Consider a pair $p', p \in (0, 1)$ that satisfies $\epsilon < p' - p \leq 2\epsilon$. If ϵ is sufficiently small, then there exists some skill interval $[w_1, w_2] \subset W$, with $\theta_H w_1 > \theta_L w_2$ and the following property: under assumption 3.5, for all p, almost surly, there exist coalitions $J_L \subset I$ and $J_H \subset I$ such that:

i) All members of J_L and J_H have a skill parameter within $[w_1, w_2]$. Moreover, for all $i \in J_L$, $\theta_i = \theta_L$ and for all $i \in J_H$, $\theta_i = \theta_H$.

ii) Both coalitions are of equal size, possess no subcoalitions and satisfy

$$p' - p \;\; = \;\; \mu(J_L) = \mu(J_H) \; .$$

Suppose first that the true average taste parameter is given by p. As $[Q, t]$ is *C-RT* there exists $i \in J_L$ such that

$$\theta_L Q(p) - \frac{K(Q(p))}{w_i} \;\; \geq \;\; \theta_L Q(p') - \frac{K(Q(p'))}{w_i} \; . \tag{3.6}$$

Suppose to the contrary that there does not exist such an $i \in J_L$. Then if all individuals in J_L announce a high taste parameter this yields a partial taste manipulation which satisfies *weak undetectability*, *unanimity*, *individual stability*, because due to the *NDT* property individuals are willing to announce any taste parameter, and *collective stability*, as J_L has no subcoalition. Now suppose that the true aggregate taste level is given by p'. Analogously, there exists $j \in J_H$ such that

$$\theta_H Q(p') - \frac{K(Q(p'))}{w_j} \;\; \geq \;\; \theta_H Q(p) - \frac{K(Q(p))}{w_j} \; . \tag{3.7}$$

Combining the inequalities (3.6) and (3.7) yields:

$$(\theta_H w_j - \theta_L w_i)(Q(p') - Q(p)) \;\; \geq \;\; 0 \; .$$

By construction, for all $w_i, w_j \in [w_1, w_2]$, $\theta_H w_j - \theta_L w_i > 0$. Hence, it has to be true that $Q(p') \geq Q(p)$.

■

Lemma 3.5 Suppose that assumption 3.6 applies. Let the minimal coalition size ϵ be close to zero. If an allocation $[Q, t]$ with equal cost sharing satisfies *C-RT*, then: for all $w \in W$, and for any pair $p', p \in (0, 1)$ with $\epsilon < p' - p \leq 2\epsilon$, $V(p', \theta_L, w) \leq V(p, \theta_L, w)$ and $V(p', \theta_H, w) \geq V(p, \theta_H, w)$.

Proof Without loss of generality, suppose that there exist p' and p with $\epsilon < p' - p \leq 2\epsilon$ and $w \in W$ such that $V(p, \theta_H, w) > V(p', \theta_H, w)$. As ϵ is small, there exists an interval $[w_1, w_2] \subset W$ with $w_1 \leq w \leq w_2$ and $w_1 < w_2$ and a coalition J_H with the following properties: For all $i \in J_H$, $w_i \in [w_1, w_2]$ and $\theta^i = \theta_H$ and, moreover, $\mu(J_H) = p' - p$. Obviously, if the true average taste parameter equals p, this coalition possesses a partial taste manipulation. This contradicts the *C-RT* property of $[Q, t]$.

■

As a consequence of Lemmas 3.4 and 3.5, whenever the provision rule Q is such that for some pair p', p with $p < p'$ one has $Q(p) > Q(p')$ or, $V(p', \theta_L, \bar{w}) > V(p, \theta_L, \bar{w})$ or $V(p', \theta_H, \underline{w}) < V(p, \theta_H, \underline{w})$, then there exists a value for the minimal coalition size ϵ such that the *C-RT* property is violated.

Lemma 3.6 If for any $w \in W$, $V(p, \theta_L, w)$ is non-increasing and $V(p, \theta_L, w)$ is non-decreasing in p then the *C-RT* property is implied, for any minimal coalition size $\epsilon > 0$.

Proof Suppose that $[Q, t]$ does not have the *C-RT* property. Then there exists a level of the true aggregate taste parameter p and a coalition J with a partial taste manipulation which induces an announced aggregate taste level of $p' \neq p$. Without loss of generality, assume that $p' > p$. Suppose that J contains an individual with a low taste parameter. Due to the *unanimity* property, this individual is made strictly better off by this partial taste manipulation. This contradicts the assumption that $V(p, \theta_L, \bar{w})$ is non-increasing in p. Now suppose that J contains only of individuals with a high taste parameter. If the true aggregate taste level is p and individuals in J

misreport their taste parameter, this cannot induce an announced aggregate taste level exceeding p.

■

Lemma 3.7 Suppose that assumption 3.6 applies. Consider an allocation $[Q, t]$ with equal cost sharing. Let Q be a non-decreasing function of p. Then, the *C-RT* property holds if $V(p, \theta_L, \bar{w})$ is non-increasing in p, and $V(p, \theta_H, \underline{w})$ is non-decreasing in p.

Proof If $V(p, \theta_L, \bar{w})$ is non-increasing in p, one has for all p and all p' with $p' \geq p$ that

$$\theta_L \bar{w}(Q(p') - Q(p)) \quad \leq \quad K(Q(p')) - K(Q(p)) \ .$$

As Q is non-decreasing in p, this implies that $\forall w \in W$,

$$\theta_L w(Q(p') - Q(p)) \quad \leq \quad K(Q(p')) - K(Q(p)) \ .$$

Hence, for all w, $V(p, \theta_L, w)$ is non-increasing in p. Analogously one shows that if $V(p, \theta_H, \underline{w})$ is non-decreasing in p, this implies that, for all w $V(p, \theta_H, w)$ is non-decreasing in p. Using Lemma 3.6 this establishes the *C-RT* property.

■

Chapter 4

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

4.1 Introduction

This paper combines the problem of optimal income taxation with the free-rider problem in public good provision. An optimal income tax is based on the utilitarian desire to redistribute resources in favor of the less able. An optimal solution of the free-rider problem has the property that a public good is installed if and only if the aggregate valuation in the economy is sufficiently high. The present paper studies the interaction between these problems. It arises because expenditures on public goods and on income transfers are linked through a public sector budget constraint. That is, they compete for the same funds. Consequently, an individual's view on the desirability of public good provision will depend on the way he is treated by the transfer system.

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

The difficulty in finding an optimal mechanism for both redistribution and public good provision is that there are two incentive problems simultaneously. The first one is familiar from the theory of optimal income taxation and is due to the fact that individuals have private information on their earning abilities. This imposes incentive constraints on redistribution which give rise to what is known as the *equity-efficiency tradeoff*.[1] The second problem is the classical *free-rider problem*, which arises because individuals have private information on their valuation of a non-excludable public good.

The main insight from the joint analysis of these two incentive problems is that the *equity-efficiency tradeoff* and the *free-rider problem* interact in a systematic fashion. More able individuals can have an excessive desire for public good provision, which they value as an instrument to limit the extent of redistribution. Likewise, less able individuals may tend to understate the desirability of provision in order to avoid a cut of transfers. Hence, a deci-

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

sion on provision that reflects the "true" aggregate valuation of the public good necessitates an adjustment of the transfer system that corrects these biases. This requires a *complementarity* between the level of redistribution and the decision on public good provision, relative to an *equity-efficiency tradeoff* without a *free-rider problem*: To prevent the more productive class from exaggerating, public good provision has to be accompanied by an increased level of redistribution. Similarly, the less productive are prevented from understating their valuation of the public good by a reduced level of redistribution if there is no public good provision.

The model that is used to arrive at these results combines a *screening problem* with a problem of *information aggregation* and involves two dimensions of individual heterogeneity, earning abilities as well as preferences for the public good. More precisely, the following assumptions are made. Individuals either have a low or a high level of earning ability.[2] Likewise, valuations of the public good are either high or low. Moreover, public goods preferences are assumed to be perfectly correlated with earning ability. That is, all individuals with the same level of earning ability also have the same valuation of the public good. With this specific information structure, the *screening problem* is to identify which individual has been assigned which level of earning ability. The problem of *information aggregation* is the elicitation of the public goods preferences of high and low ability individuals, respectively.[3]

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

[3]Consequently, the *screening* problem is based on only one dimension of individual heterogeneity. There cannot be a discrimination between individuals with the same earning ability but different public goods preferences, as in Hellwig (2004). This author however

As is standard in the literature on optimal income taxation, the present paper assumes that there is a continuum of agents. While this assumption has a variety of convenient implications, it creates a difficulty when trying to discuss problems of *information aggregation* under incentive constraints. One might argue that, in a large economy, *free-rider problems* do not arise as a single individual has no impact on public good provision and hence no reason to hide his true valuation. However, the present paper takes a different view, based on the observation that, in a continuum economy, collective behavior of individuals has an impact on the perceived aggregate valuation of the public good. Indeed as will be shown below, allocation rules based on income tax schedules are vulnerable to coordinated manipulations by large groups of agents. The notion of a *collectively incentive compatible* income tax is introduced to deal with this issue. It specifies collective incentive conditions that ensure that *information aggregation* may proceed even under the threat of manipulative collective behavior.[4]

The remainder of the paper is organized as follows. Section 4.2 defines the environment. As a benchmark, Section 4.3 derives the optimal income taxation without a *free-rider problem*. Section 4.4 contains the definition of a *collectively incentive compatible income tax*. In section 4.5 the *optimal* collectively incentive compatible income tax is characterized. The last section contains concluding remarks. All proofs can be found in the appendix.

assumes that there is no problem of *information aggregation*.

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

4.2 The environment

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

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

The random variables $\tilde{\theta}^j$, $j \in I$, are assumed to be perfectly correlated among all individuals with the same productivity parameter, i.e. ex post all individuals of class t, $t \in \{1,2\}$, have the same taste parameter. Let θ_t be the common value of the taste parameter $\tilde{\theta}^j$ for all individuals j with $\tilde{w}^j = w_t$. The taste parameters θ_1 and θ_2 are the realizations of random variables $\tilde{\theta}_1$ and $\tilde{\theta}_2$. The economy as a whole is subject to uncertainty about these random variables. There are four possibilities, or *states*, denoted by s_{LL} , s_{LH} ,s_{HL} and s_{HH}, where, e.g. s_{LL} indicates that $\tilde{\theta}_1 = \theta_L$ and $\tilde{\theta}_2 = \theta_L$. Analogously, s_{LH} indicates that $\tilde{\theta}_1 = \theta_L$ and $\tilde{\theta}_2 = \theta_H$, etc. The set of states is written as $S = \{s_{LL}, s_{LH}, s_{HL}, s_{HH}\}$.

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

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

$$U_t \;=\; \theta_t Q + u(C) - v\left(\frac{Y}{w_t}\right) . \tag{4.1}$$

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

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

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

Information Structures

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

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

Anonymous Allocations and Income Tax Mechanisms

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

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

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

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

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

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

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

If the information structure exhibits private information on taste parameters, then the tax setting institution has to deduce the actual state from individual reports. That is, the mechanism designer receives from each individual a statement which consists of an announced earning ability level and an announced taste parameter. An evaluation of all individual reports makes it possible to observe whether the less (more) able individuals have a low or a high taste parameter. The fact that s can not be taken as given explains why the message set in the revelation game equals Γ. However, if the analysis is concerned with pure assignment uncertainty, the message set $\Gamma_r(s)$ is sufficient.[6]

The *I-IC* conditions require that truth-telling constitutes an equilibrium in weakly dominant strategies.[7] That is, truth-telling has to be a best-response from an individual's perspective, irrespective of the announcements of others and irrespective of the actual state of the world. However, the *I-IC* conditions specify individual incentives only in response to message profiles that indicate a feasible state of the economy. A complete description of the revelation game also requires a specification of what happens if this distribution is incompatible with what is commonly known about the set S. These out-off-equilibrium payoffs have to preserve the incentive structure, i.e. they have to be such that truth-telling is a weakly dominant strategy. This is, for instance, achieved by choosing $Q = 0$ and a degenerate consumption-income

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

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

menu that contains only one C-Y-combination.

The final remark clarifies why the set of anonymous, feasible and *I-IC* allocation mechanisms is of relevance for an analysis of income tax systems. To this end, call an anonymous allocation mechanism an *income tax mechanism* if there exists a function $T : \mathbb{R}_+ \times S \to \mathbb{R}$ such that $\forall (w,\theta) \in \Gamma, \forall s \in S$:

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

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

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

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

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

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

Lemma 4.1 An anonymous allocation mechanism is an income tax mechanism if and only if it is feasible and possesses the following properties:

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

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

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

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

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

4.3 Pure assignment uncertainty

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

4.3.1 The optimization problem

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

Under pure assignment uncertainty, an income tax mechanism is a collection $\{Q(s), Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s \in S}$, which satisfies, for all s, the feasibility constraints

$$Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq kQ(s), \quad Q(s) \in \{0, 1\} \,, \tag{4.2}$$

and the *I-RP* constraints

$$\begin{aligned} u(C_1(s)) - v\left(\frac{Y_1(s)}{w_1}\right) &\geq u(C_2(s)) - v\left(\frac{Y_2(s)}{w_1}\right) \,, \\ u(C_2(s)) - v\left(\frac{Y_2(s)}{w_2}\right) &\geq u(C_1(s)) - v\left(\frac{Y_1(s)}{w_2}\right) \,. \end{aligned} \tag{4.3}$$

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

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

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

Under pure assignment uncertainty, the state s is commonly known. Hence, it might seem natural to define an *optimal* utilitarian income tax mechanism such that, for given s, $W(s)$ is maximized subject to the feasibility constraints in (4.2) and the *I-RP* constraints in (4.3). I will, however, proceed differently. Below a definition is stated which yields trivially the same set of optimal allocations, but facilitates a comparison to the case of private information on taste parameters discussed in later sections.

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

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

Definition 4.1 Under pure assignment uncertainty, an optimal income tax mechanism chooses $\{Q(s), Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s\in S}$ in order to maximize EW subject to the feasibility constraints in (4.2) and the *I-RP* constraints in (4.3).

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

Characterizing the informed optimum

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

$$\begin{aligned}
\max_{C_1,Y_1,C_2,Y_2} \quad & u(C_1) - v\left(\tfrac{Y_1}{w_1}\right) + u(C_2) - v\left(\tfrac{Y_2}{w_2}\right) \\
\text{s.t.} \quad & Y_1 - C_1 + Y_2 - C_2 \geq r \;, \\
& u(C_1) - v\left(\tfrac{Y_1}{w_1}\right) \geq u(C_2) - v\left(\tfrac{Y_2}{w_1}\right) \;, \\
& u(C_2) - v\left(\tfrac{Y_2}{w_2}\right) \geq u(C_1) - v\left(\tfrac{Y_1}{w_2}\right) \;.
\end{aligned} \tag{4.4}$$

A solution to problem (4.4) is parameterized by the revenue requirement r and denoted $(Y_1^*(r), C_1^*(r), Y_2^*(r), C_2^*(r))$. The following result is well known (see e.g. Stiglitz (1982)).

Lemma 4.2 At a solution to problem (4.4) the feasibility constraint and only the *I-RP* constraint for $t = 2$ are binding, implying that there is a *distortion at the bottom* and *no distortion at the top*:

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

Intuitively, problem (4.4) is essentially a problem of redistribution under incentive constraints. As the more productive suffer less from the necessity to generate income, a utilitarian planner wants them to work harder. This implies a binding *I-RP* constraint for this class of individuals at the informed optimum.

The informed optimum is now characterized with reference to problem (4.4). I use a shorthand notation for the utility level at a solution to problem (4.4) induces for type t individuals:[8]

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

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

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

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

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

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

Assumption 4.1 An informed planner chooses to install the public good in all states except state s_{LL}:[9]

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

For ease of reference, I denote by $Q^i : Q = 0 \Longleftrightarrow s = s_{LL}$ the provision rule chosen by an informed planner. To complete the description of the informed

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

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

optimum, I denote by $U_1^i(s)$ and $U_2^i(s)$ the realized utility levels of class 1 and class 2 individuals. Obviously,

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

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

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

4.3.2 Conflicting interests at the informed optimum

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

$$R_1(0) - R_1(k) \;>\; \theta_H \;>\; \theta_L \;>\; R_2(0) - R_2(k)\,. \tag{4.5}$$

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

A clarification of the possible patterns of conflicting interests will be important for an understanding of the additional incentive problems that come into play under an information structure with private information on taste parameters. Intuitively, if the scenario characterized by the inequalities in (4.5) arises, less productive individuals want to prevent the public good from being installed in every state s, and hence they have an incentive to report a low taste realization even if in fact their taste parameter is high. Likewise, the more able class wants to get the public good in every state and might be tempted to report a high taste in case of a low taste realization.

The following lemma is important for an understanding of possible scenarios of conflicting interests. It shows that for the less productive class of individuals the utility loss is larger if in problem (4.4) the revenue requirement r is increased. In more technical terms, the lemma establishes a property of *decreasing differences* according to which a lower productivity level translates into a larger utility loss. The proof relies on the following assumption:

Assumption 4.2 The function v is strictly convex and satisfies[10]

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

Lemma 4.3 Let $v(\cdot)$ satisfy Assumption 4.2. Let $r' > r$. Then:

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

[10]Note that a sufficient condition for Assumption 4.2 is $v''' \geq 0$. An alternative assumption, which would also yield the result of Lemma 4.3, is that the function v is linear. For a discussion of this quasi-linear case, see Weymark (1986) or Boadway et al. (2000).

The intuition behind this observation is as follows: Consider a solution to problem (4.4) and suppose the revenue requirement is slightly increased. The more productive cannot be forced to cover the resulting small budget deficit, as this would violate their *I-RP* constraint. To the contrary, less able individuals can be made worse off without violating any constraint. Consequently, the planner has to make them worse off if there is a need to extract larger revenues.

Possible scenarios of conflicting interests

If combined with the observation that the utility loss is larger for less able individuals, as shown in lemma 4.3, assumption 4.1 implies that the willingness of less able individuals to pay is negative if $\theta_1 = \theta_L$. Analogously, for the more productive class, the willingness to pay is positive if $\theta_2 = \theta_H$,

$$R_1(0) - R_1(k) > \theta_L \quad \text{and} \quad \theta_H > R_2(0) - R_2(k) \; . \tag{4.6}$$

These inequalities in conjunction with assumption 4.1 reduce the set of possible parameter constellations. The following three scenarios may arise.

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

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

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

These inequalities are interpreted as follows.

Scenario 1: For individuals of any class t, willingness to pay for the public good is positive if the taste realization is high, $\tilde{\theta}_t = \theta_H$, and is negative if the taste realization is low, $\tilde{\theta}_t = \theta_L$. Scenario 1 hence gives rise to the statement that, at the *informed optimum*, *willingness to pay for the public good is independent of earning ability.*

Scenario 2: For the less productive class, as under *Scenario 1*, the willingness to pay for the public good is positive only if the utility gain is high. In contrast, more productive individuals, whose utility loss is smaller, have a positive willingness to pay in any state s.

Scenario 3: For more productive individuals, as under *Scenario 2*, the willingness to pay for the public good is always positive. In addition, less able individuals suffer from such a heavy utility loss that their willingness to pay is negative in any state s.

4.4 Private information on taste parameters

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

To illustrate this, suppose *Scenario 2* applies and ask whether the informed optimum is implementable. If one takes the view that individual incentives are enough, the answer is yes. As all individuals take the mechanism designer's perception of the actual state as outside their influence, no isolated individual has a reason to misreport her own taste parameter. However, if the informed optimum is implemented, the more productive individuals want to have the public good in all states of the world, that is, even if $\tilde{\theta}_2 = \theta_L$. And moreover, if class 2 individuals are able to convince the utilitarian planner that their taste parameter is in fact high, they can ensure the provision of the public good. As the decision on provision is based on a revelation game, there is an obvious way to achieve this: a collective lie of all class 2 individuals on their taste parameter.

These considerations highlight the following issues: *First*, with private information on taste parameters, a mechanism designer may not be able to detect a deviation from the truth by a subset of agents. If all class 2 individuals make the same announcement $\hat{\theta}_2$, it is not possible to tell whether those individuals are jointly lying or are jointly telling the truth. *Second*, such a deviation may be beneficial for such a subset of agents. *Third*, it is not prevented by individual incentive compatibility. Given that all class 2 individuals lie about their taste parameter, there is no incentive for an isolated class 2 individual to reveal the realization of $\tilde{\theta}_2$ truthfully. Due to the *NDT-U* property of income tax mechanisms, this is a systematic feature. A collective deviation involving taste parameters is not undermined by individual incentives.

4.4.1 Collective Incentive Compatibility

In the following a *collectively incentive compatible (C-IC)* income tax mechanism is defined. Such a mechanism ensures that truth-telling is an equilibrium outcome even under the threat of collective manipulations.

Denote by $\mathcal{J}$ the set of measurable subsets of the set of agents, $I = [0, 1]$, with positive length. A typical element is denoted J. Denote the true profile of characteristics in J by $\gamma_J := \{(w^j, \theta^j)\}_{j \in J}$. Denote the reported profile by $\hat{\gamma}_J := \{(\hat{w}^j, \hat{\theta}^j)\}_{j \in J}$.

Denote the cross-section distribution of announcements induced by $\hat{\gamma}_J$ if the true state of the economy is $s \in S$ and all individuals not in J report truthfully by $\delta(\hat{\gamma}_J, s)$. Note that any such distribution belongs to the set $\Delta(\Gamma)$ of probability distributions on $\Gamma = \{w_1, w_2\} \times \{\theta_L, \theta_H\}$, i.e. it assigns a probability weight to each of the four elements of Γ.

Denote by $\mathcal{D} := \{d_{LL}, d_{LH}, d_{HL}, d_{HH}\}$ the set of cross-section distributions of characteristics which correspond in an obvious way to the possible states of the world, e.g. d_{LH} is a distribution which assigns equal mass to the elements of $\Gamma_r(s_{LH}) = \{(w_1, \theta_L), (w_2, \theta_H)\}$. For $\delta(\hat{\gamma}_J, s) \in \mathcal{D}$, denote by $\hat{s}(\hat{\gamma}_J, s) \in S$, the *perceived* state of the world, e.g. if $\delta(\hat{\gamma}_J, s) = d_{LH}$, then $\hat{s}(\hat{\gamma}_J, s) = s_{LH}$.

Definition 4.2 A coalition J is said to *manipulate* an income tax mechanism if there exists $s \in S$ and $\hat{\gamma}_J \neq \gamma_J$ with the following properties:

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

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

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

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

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

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

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

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

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

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

A peculiarity of Definition 4.2 is that the collective stability of a coalition J is defined with reference to the collective stability of a coalition $K \subset J$. Obviously, in a continuum economy, there is no chance of tracing these notions back to the collective stability of some "smallest" coalitions. As will become clear (see Proposition 4.1), for the purposes of this paper, this does not create a problem. The structure of a two-class economy is sufficiently simple to arrive at a complete characterization of *C-IC* income tax mechanisms.

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

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

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

Definition 4.3 A *utility allocation* specifies for every state $s \in S$, utility levels $\tilde{U}_1(s)$ and $\tilde{U}_2(s)$ for type 1 and type 2 individuals, respectively. A utility allocation is said to be *implementable* if there exists a *C-IC* income

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

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

tax mechanism such that $\forall s \in S$, and for all $(w_t, \theta_t) \in \Gamma_r(s)$,

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

where $V_t(s)$ is a shorthand for the utility class t individuals derive in state s from their consumption-income combination.

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

Proposition 4.1 Suppose there is no pooling of earning ability, that is, $\forall s \in S, \forall (w_t, \theta_t) \in \Gamma_r(s), (C(w_1, \theta_1, s), Y(w_1, \theta_1, s)) \neq (C(w_2, \theta_2, s), Y(w_2, \theta_2, s))$.[13] Then, a utility allocation is Pareto-optimal if and only if it is implementable by a feasible allocation mechanism which satisfies *I-RP* and the following properties:

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

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

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

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

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

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

The *NDT-CY* property requires that, for a given distribution of characteristics in the economy, the allocation of private goods is independent of taste parameters. This is a slightly stronger property as relative to *NDT-U*. According to the *C-RT-C*-property, manipulations of coalitions consisting only of individuals with the same type and which misreport only the taste parameter are ruled out. Obviously, this condition is necessary for *C-IC*. Proposition 4.1 states that it is also sufficient if one restricts attention to Pareto-optimal allocations.

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

Proposition 4.1 justifies the restriction to allocation rules with the *NDT-CY* property. This implies that a more concise notation can be used. In the following, the consumption and income for an individual of type t, given that the state of the world is s, is written as $(C_t(s), Y_t(s))$, with the understanding that this pair equals both $(C(w_t, \theta_L, s), Y(w_t, \theta_L, s))$ and $(C(w_t, \theta_H, s), Y(w_t, \theta_H, s))$. The *I-RP* property is hence written in the following as $\forall s \in S$, $\forall t$, $\forall t' \neq t$,

$$u(C_t(s)) - v\left(\frac{Y_t(s)}{w_t}\right) \geq u(C_{t'}(s)) - v\left(\frac{Y_{t'}(s)}{w_t}\right) . \tag{4.7}$$

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

$$Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq kQ(s) . \tag{4.8}$$

The set of implementable allocation rules is represented in the remainder of the paper by the collections $\{Q(s), Y_1(s), C_1(s), Y_2(s), C_2(s)\}_{s \in S}$, which satisfy the *C-RT-C* property, as well as the inequalities in (4.7) and (4.8). The optimal utilitarian income tax mechanism is now defined as follows.

Definition 4.4 With private information on taste parameters, the *optimal C-IC* income tax solves the problem of choosing $\{Q(s)$, $Y_1(s)$, $C_1(s)$, $Y_2(s)$, $C_2(s)\}_{s \in S}$, subject to the *C-RT-C* constraint, the *I-RP* constraints in (4.7) and the feasibility constraints in (4.8), in order to maximize EW.

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

4.5 Optimality under Collective Incentives

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

Lemma 4.4 Under *C-RT-C*, provision rules are increasing in both arguments, $\forall x \in \{L, H\} : Q(s_{xL}) \leq Q(s_{xH})$ and $\forall y \in \{L, H\} : Q(s_{Ly}) \leq Q(s_{Hy})$.

The monotonicity constraints stated in the lemma imply that there are only six candidate provision rules.[14] The provision rule $Q^i : Q = 0 \Longleftrightarrow s = s_{LL}$, which is part of the informed optimum, satisfies these constraints. The same is true for provision rule $Q^{i'}$, defined by $Q = 1 \Longleftrightarrow s = s_{HH}$, provision rule Q^1, which calls for public good provision if and only if class 1 individuals have a high taste parameter $Q^1 : Q = 1 \Longleftrightarrow s \in \{s_{HL}, s_{HH}\}$, and the analogously defined provision rule $Q^2 : Q = 1 \Longleftrightarrow s \in \{s_{LH}, s_{HH}\}$. Finally, the monotonicity constraints are trivially satisfied by the constant provision rules $Q \equiv 0$ and $Q \equiv 1$.

One of these six candidate provision rules is taken as given when undertaking the *first* step. The subsequent analysis focuses on the problem of finding an optimal *C-IC* income tax that implements the informed planner's provision rule Q^i. Formally, this problem is denoted *Problem* P^i and defined as follows.

[14] The lemma follows from standard arguments. See the appendix.

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

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

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

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

$$\begin{aligned} & V_1(s_{LH}) = V_1(s_{HH}) \ , \quad \theta_H \geq V_1(s_{LL}) - V_1(s_{HL}) \geq \theta_L \ , \\ & V_2(s_{HL}) = V_2(s_{HH}) \ , \quad \theta_H \geq V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L \ , \end{aligned} \tag{4.9}$$

the *I-RP* constraints in (4.7) and the feasibility constraints

$$\begin{aligned} & Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq 0, \quad \text{for } s = s_{LL} \ , \\ & Y_1(s) - C_1(s) + Y_2(s) - C_2(s) \geq k, \quad \text{otherwise} \ . \end{aligned} \tag{4.10}$$

4.5.1 When does collective incentive compatibility matter?

With reference to problem P^i, the *Scenarios* for which the *informed optimum* survives the introduction of collective incentive requirements are easily clarified. Recall that the informed optimum is obtained by maximizing EW_V subject to *I-RP* and feasibility, without taking *C-RT-C* into account. Obviously, the informed optimum satisfies *C-RT-C* if and only if the statements in (4.9) remain true as one replaces $V_t(s)$ by $R_t(0)$ if $s = s_{LL}$ and by $R_t(k)$ if

[15] One arrives at the inequalities in (4.9) by plugging Q^i into the *C-RT-C* constraints.

$s \neq s_{LL}$. That is, the informed optimum satisfies *C-RT-C* if and only if

$$\theta_H \geq R_1(0) - R_1(k) \geq \theta_L \quad \text{and} \quad \theta_H \geq R_2(0) - R_2(k) \geq \theta_L \ . \tag{4.11}$$

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

Proposition 4.2 The *informed optimum* has the *C-RT-C* property if and only if *Scenario 1* holds.

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

4.5.2 How to deviate from the informed optimum?

According to Proposition 4.2, under *Scenarios 2* and *3* collective incentive problems force a deviation from the informed optimum. To understand the planner's assessment of conceivable deviations, a characterization of the *I-RP* constrained Pareto-frontier in a neighborhood of the *informed optimum* is needed. To this end, the following problem is considered. Choose

C_1, Y_1, C_2, Y_2 in order to maximze $u(C_1) - v\left(\frac{Y_1}{w_1}\right)$ subject to

$$\begin{aligned} & Y_1 - C_1 + Y_2 - C_2 \geq r \quad \text{(BC)}\ , \\ & u(C_1) - v\left(\tfrac{Y_1}{w_2}\right) \leq \bar{V}_2 \qquad \text{(I-RP}_2\text{)}\ , \\ & u(C_2) - v\left(\tfrac{Y_2}{w_2}\right) = \bar{V}_2\ . \end{aligned} \tag{4.12}$$

I denote by $P(\bar{V}_2, r)$ the utility level of class 1 individuals that is induced by solution to problem (4.12).

Lemma 4.5 Let $v(\cdot)$ satisfy Assumption 4.2.

i) For all $\bar{V}_2$ and all r, Problem (4.12) has a unique solution. This solution is such that (BC) is binding and there is *no distortion at the top.*

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

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

Part ii) of Lemma 4.5 shows that there is a well defined range of parameters such that there is indeed a tradeoff between the utility of the "rich" and the utility of the "poor".[16] Moreover the *informed utilitarian optimum* does not

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

lie at the boundary of the region where the tradeoff prevails. That is, while the utilitarian planner expands redistribution up to a level that gives rise to incentive problems – recall that the *informed optimum* has a binding *I-RP* constraint for class 2 – she does not aim at the maximal level of incentive compatible redistribution.

4.5.3 Scenario 2

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

Problem P^i under Scenario 2

Under Scenario 2, *C-RT-C* of the *informed optimum* fails as the more productive want to induce public good provision even if $\theta_2 = \theta_L$, i.e. the preferences of class 2 individuals cause a violation of the *independence* condition (4.11), and one may thus think of class 2 as the *source* of collective incentive problems. Proposition 4.3 characterizes the optimal utilitarian reaction to this problem.

Proposition 4.3 Let $v(\cdot)$ satisfy Assumption 4.2. Let the parameters θ_L and θ_H be such that *Scenario 2* arises. There exists $\bar{\theta}_L$ such that if $\theta_L \leq \bar{\theta}_L$, then a solution to *Problem* P^i has the following properties:

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

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

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

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

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

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

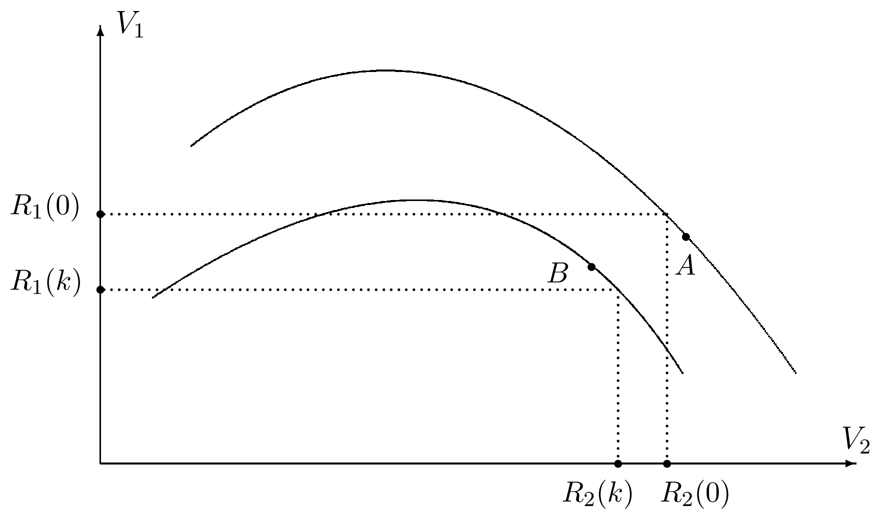

Figure 1: The graph shows the *I-RP* constrained Pareto frontiers for the revenue requirements 0 and k, respectively. Under *Sc. 2*, the difference $R_2(0) - R_2(k)$ is too small to satisfy *C-RT-C* for class 2. Under a *modest* incentive problem, the planner deviates to points A and B. Under a *severe*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

The main reason why Proposition 4.3 requires θ_L not to exceed some upper bound $\bar{\theta}_L$, is the requirement that the *C-RT-C* constraints of the less productive individuals are not binding.[17] The correction of redistribution claimed by

[17]There is also a more subtle reason. Proposition 4.3 claims that the *I-RP* constraints for class 2 are binding in all states. As is shown in the appendix, this is ensured if θ_L is sufficiently small. However, the logic of the proof does not rely on binding *I-RP* constraints of class 2 individuals, but on the shape of the Parteo frontier. As follows from Lemma 4.5,

Proposition 4.3 implies that the utility difference $V_1(s_{LL}) - V_1(s_{HL})$ shrinks relative to the *informed optimum*. If the parameter θ_L is small, then there is enough room for such an adjustment. That is, the incentive corrections required to solve P^i do not induce the less productive to prefer $Q = 1$ just in order to prevent the reduction of transfers that accompanies $Q = 0$.

These considerations suggest the following terminology for a characterization of collective incentive problems. If $\theta_L \leq \bar{\theta}_L$, incentive problems are *modest* in the sense that it is possible to correct for the "original" collective incentive problem which stems from class 2 individuals, without creating a new one resulting from class 1 individuals. By contrast, collective incentive problems are called *severe* if a solution to P^i has two binding *C-RT-C* constraints. Here, *severity* refers to the fact that the attempt to restore *C-RT-C* for one class of individuals, renders collective manipulations attractive for the other class.

Definition 4.5 Denote by $\{V_1^{**}(s)\}_{s \in S}$ the utility levels realized by class 1 individuals at a solution to problem P^i. Collective incentive problems under *Scenario 2* are called *modest* if

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

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

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

I now turn to the question whether a utilitarian planner who faces *C-RT-C* constraints indeed wants to implement provision rule Q^i. Possibly, the welfare burden of having to adjust the transfer system if Q^i is chosen is such

this shape is not affected as one enters the region where (I-RP$_2$) ceases to be binding.

that a different provision rule turns out to be superior, e.g. one alternative scheme is to install the public good in every state of the world $Q(s) = 1$ for all $s \in S$. While this provision rule has the disadvantage that resources are used to cover the cost of provision even if $s = s_{LL}$, there is no need to ask individuals about their taste parameters. Hence, there is no need to deviate from the utility levels $R_1(k)$ and $R_2(k)$, which result from the *informed optimum* if the revenue requirement equals k.

In case of a *modest* incentive problem, it depends on the planner's prior whether or not provision rule Q^i is chosen. To see this, suppose first that p_{LL} is very small. Then the provision rule $Q \equiv 1$ seems attractive, as the state in which a deviation from the *informed optimum* occurs is very unlikely, i.e. the smaller p_{LL}, the more attractive provision rule $Q \equiv 1$ becomes in comparison to Q^i. As the welfare assessment EW is continuous in the prior probabilities, there must exist prior probabilities for which $Q \equiv 1$ is superior.

Now suppose that the parameters θ_L and θ_H are such that only a "small" deviation from the informed optimum is needed to achieve collective incentive compatibility – in terms of Figure 1, the points A and B are very close to the informed optimum. In such a case, the adjustments of the transfer system, required under Q^i, are negligible in welfare terms. Consequently, one may find priors such that this provision rule remains the optimal one.

In contrast, under a *severe* incentive problem, Q^i will not be chosen. To see this, suppose that the *C-RT-C* constraints

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

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

Proposition 4.4 Let $v(\cdot)$ satisfy Assumption 4.2. Let the parameters θ_L and θ_H be such that *Scenario 2* arises.

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

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

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

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

4.5.4 Discussion of Scenario 3

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

Fortunately, these incentive corrections tend to complement each other. To see this, recall the properties of a solution to *Problem* P^i under *Scenario 2*, which was dealing only with the collective incentive problem for class 2 individuals. This solution deviates from the *I-RP* constrained Pareto frontiers for revenue requirements 0 and k, respectively, such that that the utility difference between provision and non-provision shrinks for class 1 individuals relative to the *informed optimum*, i.e. this incentive correction points in the right direction as it makes public good provision more attractive from the perspective of class 1. Hence, under *Scenario 3* the solution of *Problem* P^i may be such that the *C-RT-C* constraint for class 1 is not binding. In this case, the solution of *Problem* P^i is again characterized by Proposition 4.3.

More generally, one has to distinguish between modest and severe incentive problems. Collective incentive problems are *modest* if, at a solution to *Problem* P^i, only the *C-RT-C* constraint for one class is binding. Otherwise they are called *severe*. These collective incentive problems may imply that Q^i is not part of an optimal *C-IC* income tax mechanism.

4.6 Concluding Remarks

The analysis has shown that an optimal utilitarian income tax is robust to the introduction of a *free-rider problem* on public good provision if and only if "willingness to pay for the public good" is independent of earning ability. Otherwise, collective incentive considerations force a deviation from the optimal tax scheme. Such a deviation can take different forms, a modification of the provision rule, an adjustment of the private goods allocation accompanying a given provision rule or both. The exact pattern depends on the interaction of prior probabilities and the intensity of the collective incentive problem.

This raises the question how to assess these deviations from a welfare perspective. As the analysis has shown, it is possible that those deviations make one class better off while hurting the other class, i.e. that they do not cause a departure from constrained efficiency. However, they place an additional welfare cost on redistribution if the allocation mechanism in addition has to achieve a surplus maximizing decision on public good provision. If the latter requires that, say, the "rich" admit a low valuation of public goods, then one can not simultaneously have an excessive level of redistribution in response to such a low valuation. Consequently, one has to tradeoff the utilitarian welfare gains from a more favorable solution of the *equity-efficiency tradeoff* with those from a more favorable solution of the *free-rider problem*. This tradeoff is solved such that a deviation from an *optimal income tax*, as typically defined in the literature, is desirable in order to improve the possibility to aggregate information on the willingness to pay for public goods.

4.7 Appendix

Proof of Lemma 4.1: To proof the only if-part, note that, because preferences satisfy the separability property stated in equation (4.1), the *NDT-U*-property is an implication of *I-IC*. Obviously *I-RP* is also an implication of *I-IC*. To prove the if-part, suppose an *NDT-U* and *I-RP* allocation rule is not *I-IC*. Then there exist (w, θ) and $(\hat{w}, \hat{\theta})$ and s such that

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

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

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

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

Hence, a contradiction.

■

Proof of Lemma 4.3:

Claim 1.

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

Proof. These comparative statics are derived as follows: knowing that, at a solution to problem (4.4), the *I-RP*-constraint for type 2, as well as the budget constraint is binding allows us to setup the Lagrangean for the planner's problem. The first order conditions imply the following system of equations:

$$\frac{u'(C_1^*(r))}{u'(C_2^*(r))} = \frac{(1 - MRS_1^*) + (1 - \widehat{MRS}^*)}{MRS_1^* - \widehat{MRS}^*} , \tag{4.13}$$

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

$$Y_1^*(r) - C_1^*(r) + Y_2^*(r) - C_2^*(r) = r \ , \tag{4.14}$$

$$u'(C_2^*(r)) = \frac{1}{w_2} v'\left(\frac{Y_2^*(r)}{w_2}\right) \ , \tag{4.15}$$

$$u(C_1^*(r)) - v\left(\frac{Y_1^*(r)}{w_2}\right) = u(C_2^*(r)) - v\left(\frac{Y_2^*(r)}{w_2}\right) \ . \tag{4.16}$$

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

$$\begin{aligned} \frac{dC_2^*(r)}{dr} &= \frac{(\alpha - \delta) + \epsilon(1-\delta)}{(\alpha-\delta)(\gamma-1) + \beta(1-\delta)} \ , \\ \frac{dY_2^*(r)}{dr} &= \gamma \frac{(\alpha - \delta) + \epsilon(1-\delta)}{(\alpha-\delta)(\gamma-1) + \beta(1-\delta)} \ , \end{aligned}$$

where $\alpha :=$

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

and $\beta :=$

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

Note that, by Assumption 4.2, the common denominator of α and β is strictly positive. The numerator of β is negative because of the *distortion at the bottom* and the single crossing property, which imply that:

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

To see that the numerator of α is negative as well, note that equation (4.13) implies:

$$\begin{aligned} 2u'(C_2^*(r)) + \frac{1}{w_2} v'\left(\frac{Y_1^*(r)}{w_2}\right) - \frac{1}{w_1} v'\left(\frac{Y_1^*(r)}{w_1}\right) \ &= \\ \frac{\left[\frac{1}{w_1} v'\left(\frac{Y_1^*(r)}{w_1}\right)\right]^2 - \left[\frac{1}{w_2} v'\left(\frac{Y_1^*(r)}{w_2}\right)\right]^2}{2u'(C_1^*(r)) - \frac{1}{w_2} v'\left(\frac{Y_1^*(r)}{w_2}\right) - \frac{1}{w_1} v'\left(\frac{Y_1^*(r)}{w_1}\right)} \ &> \ 0 \ . \end{aligned}$$

Further,

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

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

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

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

Claim 2.

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

Proof. Recall that at a solution of problem (4.4), the *I-RP*-constraint of the more productive type is binding (see equation (4.16)). Hence, it must be the case that:

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

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

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

To see that also the first inequality holds, note that, using (4.15), one has:

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

This expression is strictly negative by *Claim 1.*

∎

Proof of Proposition 4.1: Before proceeding with the proof, an additional piece of notation is introduced. Recall that an income tax mechanism specifies for each s the variables $Q(s)$ and $C_1(w_1, \theta_L, s)$, $Y_1(w_1, \theta_L, s)$, $C_1(w_1, \theta_H, s)$, $Y_1(w_1, \theta_H, s)$ and $C_2(w_2, \theta_L, s)$, $Y_2(w_2, \theta_L, s)$, $C_2(w_2, \theta_H, s)$, $Y_2(w_2, \theta_H, s)$. I henceforth refer to this list of variables, as the allocation $A(s)$ for state s. An income tax mechanism M can hence be summarized as a list $M = (A(s_{LL}), A(s_{LH}), A(s_{HL}), A(s_{HH}))$. If I want to describe an income tax mechanism M' that, say, coincides with a *predefined* income tax mechanism M in all states except s_{LL} and chooses in this state the allocation prescribed by M for state s_{LH}, I write $M' = (A(s_{LH}), A(s_{LH}), A(s_{HL}), A(s_{HH}))$.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Claim 3. Consider an income tax mechanism $M = (A(s_{LL}), A(s_{LH}), A(s_{HL}), A(s_{HH}))$. Suppose there are no pooling outcomes. Suppose the induced utility allocation $\{\tilde{U}_1(s), \tilde{U}_2(s)\}_{s\in S}$ has the *C-RT-C* property and is Pareto-optimal within the set of utility allocations which are implementable by an income tax mechanism with the *C-RT-C* property. Then, there is no manipulating coalition that contains individuals of both types.

Proof. By *Claim 1*, individuals of both types reveal their productivity parameters. By non-detectability, all individuals who announce the same productivity parameter also have to announce the same taste parameter. Hence, the only conceivable manipulation that contains individuals of both types is such that all type 1 individuals misreport θ_1 and all type 2 individuals misreport θ_2. I show in the following, that Pareto-optimality within the set of *C-RT-C* utility allocations implies that there does not exist a stable joint manipulation of taste parameters that makes both type 1 and type 2 individuals strictly better off.

The proof proceeds by contradiction. Suppose there is a a joint lie on taste parameters that makes both type 1 and type 2 individuals strictly better off. Without loss of generality, assume that the true state is s_{LL}.[18] An undetectable joint collective lie induces the state perception s_{HH}. Such a collective lie makes all coalition members better off only if

$$\begin{aligned} \theta_L Q(s_{LL}) + V_1(s_{LL}) &< \theta_L Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_L Q(s_{LL}) + V_2(s_{LL}) &< \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{4.17}$$

Due to *NDT-U*, this collective deviation is individually stable, as it does only

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

involve misreports of taste parameters. To achieve collective stability as well, the following *C-RT-C* constraints have to be binding,

$$\begin{aligned} \theta_L Q(s_{LH}) + V_1(s_{LH}) &= \theta_L Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_L Q(s_{HL}) + V_2(s_{HL}) &= \theta_L Q(s_{HH}) + V_2(s_{HH}) \ ; \end{aligned} \tag{4.18}$$

otherwise a coalition consisting only of type 1 individuals or a coalition consisting only of type 2 individuals would want to deviate once more after the state perception s_{HH} has been induced.

(a) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &\geq \theta_H Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &\geq \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{4.19}$$

By (4.17), the following income tax mechanism $M' = (A(s_{HH}), A(s_{LH}), A(s_{HL}), A(s_{HH}))$ is Pareto superior. It is easily verified that M' satisfies all *C-RT-C* constraints if (4.19) holds.

(b) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &< \theta_H Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &< \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{4.20}$$

By (4.17), (4.18) and (4.20), the following income tax mechanism $M' = (A(s_{HH}), A(s_{HH}), A(s_{HH}), A(s_{HH}))$ is Pareto superior. Obviously, M' satisfies all *C-RT-C* constraints.

(c) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &< \theta_H Q(s_{HH}) + V_1(s_{HH}) \ , \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &\geq \theta_L Q(s_{HH}) + V_2(s_{HH}) \ . \end{aligned} \tag{4.21}$$

The following income tax mechanism $M' = (A(s_{HH}), A(s_{LH}), A(s_{HH}), A(s_{HH}))$ is Pareto superior. It follows from (4.17) that both types are better off in state s_{LL}. It follows from (4.18) that type 2 is not worse off in state s_{HL}, and it follows from (4.21) that type 1 is strictly better off in state s_{HL}. Moreover, it is easily verified that M' satisfies all *C-RT-C* constraints.

(d) Suppose that

$$\begin{aligned} \theta_H Q(s_{HL}) + V_1(s_{HL}) &\geq \theta_H Q(s_{HH}) + V_1(s_{HH}) \;, \\ \theta_H Q(s_{LH}) + V_2(s_{LH}) &< \theta_L Q(s_{HH}) + V_2(s_{HH}) \;. \end{aligned} \tag{4.22}$$

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

Claims 1 to 3 are summarized as follows: Suppose there are no pooling outcomes. Then a utility allocation that is Pareto-optimal within the set of utility allocations which are implementable by an income tax mechanism with the *C-RT-C* property also possesses the (more demanding) *C-IC* property. We have thus shown that the set of Pareto-optimal utility allocations under *C-IC* income tax mechanisms coincides with the set of Pareto-optimal utility allocations under *C-RT-C* income tax mechanisms. In the following, I may hence focus on the latter set. Recalling Lemma 4.1, these utility allocations are achievable by means of a feasible allocation mechanism with the *I-RP*, the *NDT-U* and the *C-RT-C* property. The following claim remains to be established:

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

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

■

Proof of Lemma 4.4: Consider for example the *C-RT-C* constraints for class 1 given that $\theta_2 = \theta_L$:

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

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

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

■

Proof of Lemma 4.5: The argument is only sketched. Consider the Lagrangean of problem (4.12):

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

Deriving first order conditions, one easily verifies that *(BC)* has to be binding, that there is *no distortion at the top* and a *distortion at the bottom* if and and only if (I-RP$_2$) 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 (4.12). Uniqueness of this solution can be established as follows. Strict quasiconcavity of preferences and the property of *no distortion at the top* uniquely determine $\bar{Y}_2$ and $\bar{C}_2$ as a function of $\bar{V}_2$. The fact that *(BC)* is binding yields a unique iso-tax-revenue line $Y_1 - C_1 = \gamma$, with $\gamma = C_2(\bar{V}_2) - Y_2(\bar{V}_2) + r$, on which the point $(\bar{Y}_1(\bar{V}_2), \bar{C}_1(\bar{V}_2))$ can be found. More precisely, $(\bar{Y}_1(\bar{V}_2), \bar{C}_1(\bar{V}_2))$ maximizes $u(C_1) - v(Y_1/w_1)$ subject to I-RP$_2$ and $Y_1 - C_1 = \gamma$. Again, due to strict quasiconcavity, the latter problem has a unique solution.

Denote the optimal values of the multipliers at the solution of problem (4.12) by $\bar{\lambda}(\bar{V}_2)$ and $\bar{\nu}(\bar{V}_2)$. These multipliers are used to study how P depends on $\bar{V}_2$. The following property is used:[19]

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

Similarly as for the proof of Lemma 4.3, comparative statics of the solution of problem (4.12) with respect to $\bar{V}_2$ can be derived. Based on this exercise, the comparative statics of the Lagrangean multipliers can be determined.[20]

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

[20] The first order conditions imply:

$$\bar{\lambda}(\bar{V}_2) = \frac{1 - \overline{MRS}_1}{1 - \widehat{MRS}} \quad \text{and} \quad \bar{\nu}(\bar{V}_2) = \frac{u'(\bar{C}_1)}{u'(\bar{C}_2)} \frac{\overline{MRS}_1 - \widehat{MRS}}{1 - \widehat{MRS}} \; , \tag{4.23}$$

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

The details of the computations are omitted. One arrives at the following results:

(a) Suppose first that (I-RP$_2$) is binding.[21] Using Assumption 4.2, 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 (I-RP$_2$) is binding the function P is strictly concave in $\bar{V}_2$ and one has:

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

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

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

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

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

$$\begin{aligned} \max_{C_1,Y_1,C_2,Y_2} \quad & u(C_1) - v\left(\tfrac{Y_1}{w_1}\right) \\ \text{s.t.} \quad & Y_1 - C_1 + Y_2 - C_2 \geq k \; , \\ & u(C_2) - v\left(\tfrac{Y_2}{w_2}\right) \geq u(C_1) - v\left(\tfrac{Y_1}{w_2}\right) \; . \end{aligned} \tag{4.24}$$

[21] The existence of a value $\bar{V}_2$ such that (I-RP$_2$) is binding follows from Lemma 4.2.

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

Denote by $\tilde{V}_2$ the utility level that results for type 2 individuals at a solution to problem (4.24). Using the first order conditions of problem (4.24) allows us to verify that $\bar{\lambda}(\tilde{V}_2) - \bar{\nu}(\tilde{V}_2) = 0$.

(e) Finally, use the first order condition (4.13) of the *informed problem* to substitute for $u'(\bar{C}_1)/u'(\bar{C}_2)$ in the formula for $\bar{\nu}(\bar{V}_2)$ (see (4.23)), one gets $\bar{\lambda}(R_2(r)) - \bar{\nu}(R_2(r)) = -1$.

■

Proof of Proposition 4.3: I consider a relaxed version of *Problem* P^i, referred to as *Problem* P^i_x. P^i_x takes only a subset of the constraints of P^i into account. I show below that a solution to P^i_x satisfies these neglected constraints.

Formally P^i_x is defined as follows. Maximize EW_V subject to the feasibility constraints in (4.10), the *I-RP* constraints for type 2 and the following subset of the *C-RT-C* constraints:

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

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

$\{V^{**}_{t,x}(s)\}_{s\in S}$ denotes the utility levels realized by class t at a solution to P^i_x. The following assumption formalizes the statement in Proposition 4.3 that θ_L must not exceed some upper bound $\bar{\theta}_L$.

Assumption 4.3 $\theta_L < \min\{\hat{R}_2(0) - R_2(k), P(R_2(k) + \theta_L, 0) - P(R_2(0) - \theta_L, r)\}$.

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

Under Assumption 4.3, Proposition 4.3 follows from the following observations.

(a) In every state s, the budget constraint is binding and there is *no distortion at the top*.

Proof. This follows from setting up the Lagrangean of P_x^i and deriving first order conditions.

(b) $V_{2,x}^{**}(s_{LL}) \leq \theta_L + R_2(k)$.

Proof. Suppose, to the contrary, that $V_{2,x}^{**}(s_{LL}) > \theta_L + R_2(k)$. Then the planner could choose, instead $(Y_1^*(k), C_1^*(k), Y_2^*(k), C_2^*(k))$ for $s \neq s_{LL}$. For $s = s_{LL}$, the planner could choose $V_{2,x}^{**}(s_{LL}) = \theta_L + R_2(k)$ and $V_1 = P(\theta_L + R_2(k), 0)$. Due to the monotonicity properties established in Lemma 4.5, this would increase utilitarian welfare in every state s.

(c) $V_{2,x}^{**}(s_{LH}) \leq R_2(k)$.

Proof. The *C-RT-C* constraints imposed under P_x^i imply $V_{2,x}^{**}(s_{LH}) \leq V_{2,x}^{**}(s_{LL}) - \theta_L$. Combing this with (b) yields (c).

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

Proof. $V_{1,x}^{**}(s_{LL}) \neq P(V_{2,x}^{**}(s_{LL}), 0)$ or $V_{1,x}^{**}(s_{HL}) \neq P(V_{2,x}^{**}(s_{HL}), k)$ immediately yields a contradiction to optimality. The *distortion at the*

[23] Recall that under *Scenario 2*, $\theta_L > R_2(0) - R_2(k)$. If θ_L does not exceed $R_2(0) - R_2(k)$ by too much, i.e. $\theta_L \simeq R_2(0) - R_2(k)$, then, the assumption $\theta_L < P(R_2(k)+\theta_L, 0) - P(R_2(0)-\theta_L, r)$ is satisfied. To see this: if $\theta_L \simeq R_2(0) - R_2(k)$, the continuity property established in Lemma 4.5 implies that $P(R_2(k)+\theta_L,0) - P(R_2(0)-\theta_L,r) \simeq R_1(0) - R_1(k)$. By definition of *Scenario 2*, the latter term exceeds θ_L.

bottom in state s_{LL} follows from Assumption 4.3 and observation (b), which imply that $V^{**}_{2,x}(s_{LL}) < \hat{R}_2(0)$.

(e) $V^{**}_{2,x}(s_{LL}) \geq R_2(0)$ and $V^{**}_{1,x}(s_{LL}) \leq R_1(0)$.

Proof. Given (d), if (e) is false, then $V^{**}_{2,x}(s_{LL}) < R_2(0)$ and $V^{**}_{1,x}(s_{LL}) > R_1(0)$. Then, using the monotonicity properties established in Lemma 4.5, it is possible to increase $V_2(s_{LL})$ and to decrease $V_1(s_{LL})$ along the (I-RP$_2$) constrained Pareto frontier without violating the constraint $V_2(s_{LL}) - V_2(s_{LH}) \geq \theta_L$, thereby increasing utilitarian welfare in state s_{LL}.

(f) $V^{**}_{2,x}(s_{LH}) = V^{**}_{2,x}(s_{HH}) = V^{**}_{2,x}(s_{HL}) =: V^{**}_{2,x}(s_H) \leq R_2(k)$.

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

Choose $(C_1(s_{LH}), Y_1(s_{LH}), C_2(s_{LH}), Y_2(s_{LH}))$ and $(C_1(s_{HH}),\, Y_1(s_{HH}),\, C_2(s_{HH}),\, Y_2(s_{HH}))$ in order to maximize

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

subject to the following constraints: Feasibility,

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

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

the *I-RP* constraints for type 2,

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

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

the *C-RT-C* constraint,

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

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

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

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

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

(g) $V^{**}_{1,x}(s_{LH}) = V^{**}_{1,x}(s_{HH}) = V^{**}_{1,x}(s_{HL}) =: V^{**}_{1,x}(s_H) = P(V^{**}_{2,x}(s_H), k) \geq R_1(k)$, implying a *distortion at the bottom* in states s_{LH}, s_{HL} and s_{HH}. Moreover, at a solution to *Problem* P^i_x $(C_1(s_{HL}), Y_1(s_{HL})) = (C_1(s_{LH}), Y_1(s_{LH})) = (C_1(s_{HH}), Y_1(s_{HH})) =: (C_1(s_H), Y_1(s_H))$ and $(C_2(s_{HL}), Y_1(s_{HL})) = (C_2(s_{LH}), Y_1(s_{LH})) = (C_2(s_{HH}), Y_1(s_{HH})) =: (C_2(s_H), Y_2(s_H))$.

Proof. The first statement follows from (d), (f), optimality considerations and the monotonicity properties established in Lemma 4.5, making use of the fact that $V^{**}_{2,x}(s_H) \leq R_2(k)$. The equality of (C_t, Y_t)

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

pairs across states follows from the uniqueness established in Lemma 4.5.

(h) $V^{**}_{1,x}(s_H) \neq R_1(k)$ and $V^{**}_{2,x}(s_H) \neq R_2(k)$ and $V^{**}_{2,x}(s_{LL}) \neq R_2(0)$ and $V^{**}_{1,x}(s_{LL}) \neq R_1(0)$ and $V^{**}_{2,x}(s_{LL}) - V^{**}_{2,x}(s_{LH}) = \theta_L$.

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

(i) $\theta_H > R_1(0) - R_1(k) > V^{**}_{1,x}(s_{LL}) - V^{**}_{1,x}(s_H)$

Proof. This follows from (e) (g), (h) and the definition of *Scenario 2*.

(j) $V^{**}_{1,x}(s_{LL}) \geq P(R_2(k) + \theta_L, 0)$ and $V^{**}_{1,x}(s_H) \leq P(R_2(0) - \theta_L, r)$.

Proof. The first inequality follows from (b), (d) and the monotonicity property established in Lemma 4.5. The second inequality is established as follows: Analogously as in (b), one shows that $V^{**}_{2,x}(s_H) \geq R_2(0) - \theta_L$ and then uses $V^{**}_{2,x}(s_H) = P(V^{**}_{2,x}(s_H), k)$ and again the monotonicity property.

(k) At solution to P^i_x, the neglected *C-RT-C* constraint for class 1 is satisfied. I.e. $\theta_H > V^{**}_{1,x}(s_{LL}) - V^{**}_{1,x}(s_H) > \theta_L$

Proof. This follows from (i), (j) and Assumption (4.3).

■

Chapter 5

Distortionary Taxation and the Free-Rider Problem

5.1 Introduction

This paper derives the optimal utilitarian rule for public good provision under the premise that the costs are covered via distortionary taxation *and* that individuals have private information of their valuation of the public good. The existing literature has been concerned *either* with the impact of distortionary taxation *or* with the consequences of private information.

One branch of the literature is the theory of optimal taxation. It characterizes the optimal quantity of a public good by a *modified* Samuelson rule that equates the marginal costs of public funds and the sum of marginal utilities.[1] In this framework, a problem of preference elicitation does not arise because the distribution of preferences in the economy is assumed to be common knowledge.

[1]Examples include Atkinson and Stern (1974), Wilson (1991), Boadway and Keen (1993), Nava et al. (1996), Sandmo (1998), Hellwig (2004) and Gaube (2000, 2005).

The second branch is the literature on the *free-rider problem* in public good provision, which studies the question what an optimal allocation of public goods looks like if individuals have private information of their preferences of a public good.[2] This literature focusses on quasi-linear environments in which the marginal disutility of having to pay for the public good is constant for all individuals. Hence, in this approach, payment obligations do not drive a wedge between marginal rates of substitution and marginal rates of transformation; that is, taxation is not distortionary.

The contribution of this paper is the derivation of a *twice modified* Samuelson rule that takes into account the welfare burden of distortionary taxes and, in addition, the desirability of preference revelation. This concern is not only driven by the desire to have a more complete theory. There is an interesting economic relationship between these two issues: Whether or not an individual is willing to reveal her valuation of a public good depends on the way she is treated by the tax system.

To see this, suppose that, as in this paper, a linear tax on income is used for public goods finance. Consequently, individuals with a higher level of income pay more taxes and hence contribute more to the cost of public good provision. When asked to report their preferences, these individuals compare their utility gain from public good provision to their additional tax burden. The fact that they have to contribute relatively large payments may imply that they refuse to reveal their true valuation of the public good. Instead, they choose their announcement such that they prevent the public good from being provided.

[2] The seminal contributions in the early literature are Clarke (1971), Groves (1973) and Green and Laffont (1977). See Hellwig (2003) or Norman (2004) for more recent treatments.

In more general terms, the tax system affects the willingness of individuals to reveal their valuation of a public good and is thus a potential source of incentive problems. This paper is a first attempt to discuss what an optimal response to these incentive problems looks like in a model which is as simple as possible: Individuals either have a low or a high level of earning ability. Likewise, valuations for the public good are either high or low. For the economy as a whole, there is uncertainty about the public goods preferences of the "rich" class and the "poor" class of agents, respectively. The decision on public good provision is binary, i.e. the public good is either installed or not.[3] Finally, the tax instrument used to finance public good provision is a linear tax on income.

An analysis of the *free-rider problem* under distortionary taxation can be conducted for all kinds of tax instruments, an affine linear income tax, a non-linear income tax, a combination of direct and indirect tax instruments, etc. This paper focusses on a liner tax on income that is raised only to cover the cost of public good provision. This setup has the advantage of simplicity. In particular, it is easily seen how the tax system shapes individual assessments of the public good: The formal analysis proceeds under the assumption that individuals reduce their labor supply in response to an increased tax on income. Under this premise, it can be shown that individuals with a high level of earning ability suffer ceteris paribus from a larger utility loss if additional taxes are raised. Consequently the burden of taxation for a public good that is enjoyed by individuals of both classes is essentially carried by the "rich" class.

This generates the following pattern of incentive problems: More able indi-

[3] Chapter 2 do not assume a perfect correlation of earning ability and public goods preferences, and they allow for a continuum of different provision levels.

viduals tend to understate their willingness to pay for the public good because they suffer more intensively from an increase of the tax revenue requirement. Analogously, less able individuals exaggerate when asked about their valuation because they don't feel a large utility burden from higher taxes.

As an example, think of the decision whether or not to use public money for the construction of a park. For the sake of the argument, suppose that the true state of the world is such that all inhabitants of the town realize the same utility gain if the park is available. Due to the tax system, however, the decision whether or not to install it creates conflicting interests between individuals with a high level of income and individuals with a low level of income. These interests govern the behavior of individuals in the revelation game and thus create an impediment for the acquisition of information on preferences.

The example illustrates that a pattern of incentive problems where less able individuals tend to exaggerate their preferences and the more able are too reserved is a plausible case.[4] As a consequence, incentive compatibility constraints imply that the *twice modified* Samuelson rule relies on the use of excessive taxes, i.e. of taxes which are larger than actually needed to cover the cost of public good provision. These are used for two different reasons. Either they serve to make public good provision artificially expensive. This case arises if incentive compatibility conditions are needed to prevent less able individuals from exaggerating their valuation of the public good; that

[4]It is, however, not the only conceivable constellation. The pattern is reversed under a non-linear income tax system, as shown in Chapter 4. If the tax system generates, in addition, direct income transfers from "rich" to "poor" households, then less able individuals oppose public good provision more intensively. Tax revenues that are spent on public goods are not available for redistribution any more. This observation is assessed differently by the "rich" and the "poor".

is, excessive taxes are used to make the public good less attractive for the less able class. Alternatively, if the more able individuals' temptation to understate their preferences causes an incentive problem, then excessive taxes can be used to make the non-provision of the public good less attractive.

If these excessive taxes become very high, then an optimal provision rule does not incorporate all pieces of information. Suppose for instance, that one needs to accompany public good provision with very high taxes in order to ensure a truthful statement from less able individuals on their valuation of the public good. Then, an optimal provision rule does not try to acquire information from them. Put differently, information that is too costly to obtain is neglected by the *twice modified* Samuelson rule.

The remainder is organized as follows. The next section specifies the model. In Section 5.3, as a benchmark, the *modified* Samuelson rule is derived. Section 5.4 solves for the *twice modified* Samuelson rule. In Section 5.5, I discuss how the possibility of direct income transfers would affect the results. The last section contains concluding remarks. All proofs are in the appendix.

5.2 The environment

The economy consists of two classes of agents that are characterized by the earning ability levels w_1 and w_2, where $w_2 > w_1$, i.e. class 2 agents are more productive. It is commonly known that there are equal shares of more and less productive individuals in the population. Individuals of class t, $t \in \{1, 2\}$, have a common taste parameter θ_t, which affects their valuation of a public good. Moreover, θ_t is private information of individuals who belong to class t.

The economy as a whole is subject to uncertainty about these taste parameters; θ_1 and θ_2 are taken to be the realizations of random variables $\tilde{\theta}_1$ and $\tilde{\theta}_2$. Both random variables can only take two values, θ_L and θ_H, where $\theta_L < \theta_H$. Consequently, there are four possible *states* of the economy, depending on the preference parameters of class 1 and class 2 individuals, i.e. depending on the actual value of the vector (θ_1, θ_2), where

$$(\theta_1, \theta_2) \in \{(\theta_L, \theta_L), (\theta_L, \theta_H), (\theta_H, \theta_L), (\theta_H, \theta_H)\} .$$

The utility function of individuals who belong to class t is given by

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

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

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

I use a mechanism design approach to characterize admissible schemes of taxation and public good provision. An *allocation rule* specifies for each state $(\theta_1, \theta_2) \in \{\theta_L, \theta_H\}^2$ a decision on public good provision $Q(\theta_1, \theta_2)$ and a linear income tax rate $\tau(\theta_1, \theta_2)$. The revenues generated by this tax are used only to cover the cost of public good provision.[5]

[5] While this is the easiest way to introduce a distortionary tax instrument into the analysis, it is certainly not the only case of interest. The impact of this assumption and

An allocation rule has to satisfy a *budget constraint (BC)*. In every state (θ_1, θ_2) the tax revenues from linear income taxation have to be sufficient to cover the cost k of public good provision. Formally, for all (θ_1, θ_2),

$$\tau(\theta_1, \theta_2)[Y_1(\tau(\theta_1, \theta_2)) + Y_2(\tau(\theta_1, \theta_2))] \ \geq \ kQ(\theta_1, \theta_2) \ , \tag{5.1}$$

where, for each $t \in \{1, 2\}$, $Y_t(\tau(\theta_1, \theta_2))$ is the utility maximizing level of effective labor supply for an individual who belongs to class t. More precisely, $Y_t(\tau(\theta_1, \theta_2))$ is the unique solution of the following maximization problem:

$$\max_Y \quad u((1 - \tau(\theta_1, \theta_2)Y)) - v\left(\frac{Y}{w_t}\right) \ . \tag{5.2}$$

The above budget constraint allows for a budget surplus, i.e. for tax rates which are higher than actually needed. It will become clear that, for incentive reasons, a deviation from budget balance may be desirable.

In addition to the budget constraint, an allocation rule has to satisfy *incentive compatibility constraints (IC)*. These constraints ensure that individuals of either class are willing to reveal their taste parameter truthfully.

$$\begin{aligned} &\forall \theta_1, \forall \hat{\theta}_1, \forall \theta_2 : \theta_1 Q(\theta_1, \theta_2) + V_1(\theta_1, \theta_2) \geq \theta_1 Q(\hat{\theta}_1, \theta_2) + V_1(\hat{\theta}_1, \theta_2) \ , \\ &\forall \theta_2, \forall \hat{\theta}_2, \forall \theta_1 : \theta_2 Q(\theta_1, \theta_2) + V_2(\theta_1, \theta_2) \geq \theta_2 Q(\theta_1, \hat{\theta}_2) + V_2(\theta_1, \hat{\theta}_2) \ , \end{aligned} \tag{5.3}$$

where $V_t(\theta_1, \theta_2)$ denotes the indirect utility function of problem (5.2),

$$V_t(\theta_1, \theta_2) \ := \ u((1 - \tau(\theta_1, \theta_2))Y_t(\tau(\theta_1, \theta_2))) - v\left(\frac{Y_t(\tau(\theta_1, \theta_2))}{w_t}\right) \ .$$

These incentive constraints are based on the presumption that, in the underlying revelation game, all individuals who belong to the same class make the same taste announcement. Under this presumption, the above inequalities ensure that the individuals of class t are not better off under a joint collective

alternative specifications of the tax system are discussed in more detail in Section 5.5.

lie about their taste parameter, whatever the collective taste announcement of individuals who belong to class $t' \neq t$. Put differently, from the class perspective the truth is required to be a dominant strategy.[6]

A more extensive discussion of these incentive constraints can be found in Chapter 3 and in Chapter 4. These papers develop the notion of a *collectively incentive compatible* income tax mechanism. The collective incentive requirement addresses the following situation: Suppose that in order to figure out the actual state of the economy, a tax setting planner has to acquire information on individual valuations of a public good. Individuals may form coalitions in order to manipulate jointly the planner's perception of the state of the economy. As a consequence, the planner will discover the true state only if he decides on public good provision in a way which eliminates all incentives for manipulative collective actions. Obviously, the above incentive constraints, which address collective actions on the class level only, are a necessary condition for *collective incentive compatibility.* For a more precise statement of conditions under which this property is also sufficient, the reader is referred to Chapter 3 and Chapter 4. For the purpose of the present paper, I just note that those conditions are met.

An *optimal* allocation rule maximizes utilitarian welfare from an ex-ante perspective, defined as a hypothetical situation where the actual state of the economy (θ_x, θ_y) is not yet known, where $x, y \in \{L, H\}$ indicate the taste realizations of class 1 and class 2 individuals, respectively. The objective

[6]The main advantage of *implementation in dominant strategies* is that the set of admissible allocations does not depend on assumptions about the prior beliefs of individuals. See e.g. Bergemann and Morris (2005); Chung and Ely (2004) or Kalai (2004).

function is a weighted average of the welfare levels (W_{xy}), where

$$W_{xy} \quad := \quad (\theta_x + \theta_y) Q(\theta_x, \theta_y) + V_1(\theta_x, \theta_y) + V_2(\theta_x, \theta_y) \ .$$

The probability weights are taken to be the prior beliefs of the tax setting planner, which are denoted $p := (p_{LL}, p_{LH}, p_{HL}, p_{HH})$, where $p_{LL} :=$ prob(θ_L, θ_L), $p_{LH} :=$ prob(θ_L, θ_H), etc. Expected welfare from the planner's ex ante perspective is accordingly given by

$$EW \quad := \quad p_{LL} W_{LL} + p_{LH} W_{LH} + p_{HL} W_{HL} + p_{HH} W_{HH} \ .$$

Definition 5.1 The optimal utilitarian allocation rule solves the problem of choosing the functions $Q : (\theta_1, \theta_2) \mapsto Q(\theta_1, \theta_2)$ and $\tau : (\theta_1, \theta_2) \mapsto \tau(\theta_1, \theta_2)$ in order to maximize EW, subject to the budget constraints in (5.1) and the incentive compatibility constraints in (5.3).

5.3 The complete information benchmark

To understand the impact of the incentive compatibility conditions, this section discusses, as a benchmark, the allocation rule that would be chosen by an *informed* utilitarian planner, i.e. a planner who happens to know the actual value of (θ_1, θ_2). For brevity, I refer to this outcome as the *informed optimum*.

Some more pieces of notation are helpful. Denote by $U^*(\tau, w_t)$ the indirect utility that is derived from consumption of private goods by an individual with earning ability w_t who faces a linear income tax rate of τ,

$$U^*(\tau, w_t) \quad := \quad \max_Y \quad u((1-\tau)Y) - v\left(\frac{Y}{w}\right) \ .$$

Denote by τ_k the linear tax rate which ensures cost coverage in case of public good provision. This tax rate is implicitly defined by the following equation,

$$\tau_k(Y_1(\tau_k) + Y_2(\tau_k)) \quad = \quad k \; .$$

Let ΔU_t^* denote the *private utility loss* of a class t individual as the tax rate increases from 0 to τ_k,

$$\Delta U_t^* \quad := \quad U^*(0, w_t) - U^*(\tau_k, w_t) \; .$$

Denote by ΔW_p^* the *aggregate private utility loss*, $\Delta W_p^* := \Delta U_1^* + \Delta U_2^*$.

The informed optimum consists of a tax rule τ^* and a provision rule Q^*. The tax rule τ^* ensures a binding budget constraint,

$$\tau^*(\theta_1, \theta_2) \quad = \quad \begin{cases} 0, & \text{if} \quad Q^* = 0 \\ \tau_k, & \text{if} \quad Q^* = 1. \end{cases}$$

Moreover, the informed planner chooses provision rule Q^* such that the public good is installed as soon as the aggregate utility gain exceeds the aggregate private utility loss from higher tax rates,

$$Q^*(\theta_1, \theta_2) \quad = \quad \begin{cases} 0, & \text{if} \quad \Delta W_p^* > \theta_1 + \theta_2 \\ 1 & \quad\;\;\text{otherwise.} \end{cases}$$

The *informed optimum* is completely characterized by the tax rule τ^* and the provision rule Q^*. However, the model outlined so far allows for a variety of different parameter constellations. For instance, if $\Delta W_p^* < 2\theta_L$, then Q^* is such that the public good is provided in every state of the world, i.e. $Q^*(\theta_1, \theta_2) = 1$ for all (θ_1, θ_2). To avoid a lengthy discussion of each conceivable parameter constellation, I focus on a particular case.

Assumption 5.1 An informed planner chooses to install the public good in all states except state (θ_L, θ_L):[7]

$$\theta_H + \theta_L \;\; \geq \;\; \Delta W_p^* \;\; \geq \;\; 2\theta_L \; .$$

For ease of reference, I denote by $Q^i : Q = 0 \iff (\theta_1, \theta_2) = (\theta_L, \theta_L)$, the provision rule chosen by an informed planner.

Conflicting interests induced by the *informed optimum*

The informed optimum may give rise to conflicting views on the desirability of public good provision. For the sake of concreteness, suppose that

$$\Delta U_2^* \;\; > \;\; \theta_H \;\; > \;\; \theta_L \;\; > \;\; \Delta U_1^* \; . \tag{5.4}$$

In this scenario, for less productive individuals, the private utility loss is so small that, in all states, they are better off if the public good is installed. By contrast, the more productive suffer so heavily as the tax rate increases from 0 to τ_k that they are always worse off if the public good is installed.

A clarification of the possible patterns of conflicting interests is important for an understanding of the impact of incentive compatibility constraints. Intuitively, if the scenario characterized by the inequalities in (5.4) arises, more productive individuals want to prevent the public good from being installed in every state and hence have an incentive to report a low taste realization, even if in fact their taste parameter is high. Likewise, the less able class wants to ensure provision and is tempted to report a high taste parameter

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

in case of a low taste realization.

The following lemma is the key to an understanding of possible scenarios of conflicting interests. It shows that, for the more productive class of individuals, the private utility loss is larger if the tax rate τ goes up. In more technical terms, the lemma establishes a property of *increasing differences* according to which a larger productivity level translates into a larger private utility loss. The proof relies on the assumption that individuals decrease their labor supply in response to an increase in the tax rate.

Assumption 5.2 Labor supply is a decreasing function of τ:[8]

$$\forall t \in \{1, 2\}, \forall \tau \in [0, 1[: \quad Y_t'(\tau) < 0 \; .$$

Lemma 5.1 For any pair of tax rates $\underline{\tau}$ and $\bar{\tau}$ with $\bar{\tau} > \underline{\tau}$. If Assumption 5.2 holds, then:

$$U^*(\underline{\tau}, w_1) - U^*(\bar{\tau}, w_1) \quad < \quad U^*(\underline{\tau}, w_2) - U^*(\bar{\tau}, w_2) \; .$$

It is easily verified that individuals with a high earning ability choose, for any tax rate, a higher level of effective labor supply; that is, the more productive class has a higher level of income and thus pays more taxes. According to Lemma 5.1, this implies that ceteris paribus it is harder to convince the more productive class of individuals that the utility gain from public good provision justifies an increase of the tax rate.[9]

[8]This assumption has been introduced by Sheshinski (1972) in a model of optimal linear income taxation and lump sum redistribution. Its role is further discussed in Hellwig (1986). An alternative assumption, which also yields the result of Lemma 5.1, is made in Persson and Tabellini (2000, Ch. 3). There income is exogenous, and utility is quasi-linear in consumption.

[9]However, the proof relies on Assumption 5.2 according to which the substitution effect associated with a higher tax rate dominates the income effect. If this relation was

If combined with Lemma 5.1, Assumption 5.1 implies that the more able class is made worse off by public good provision in case of a low taste realization, i.e. if $\theta_2 = \theta_L$. Likewise, less able individuals are better off if the public good is installed when $\theta_1 = \theta_H$; that is,

$$\Delta U_2^* > \theta_L \quad \text{and} \quad \theta_H > \Delta U_1^* \, .$$

These inequalities in conjunction with Assumption 5.1 reduce the set of possible parameter constellations. The following three scenarios may arise.

$$\begin{aligned}
&\text{Scenario 1:} \quad \theta_H \;\geq\; \Delta U_2^* \;>\; \Delta U_1^* \;\geq\; \theta_L \, , \\
&\text{Scenario 2:} \quad \theta_H \;\geq\; \Delta U_2^* \;\geq\; \theta_L \;>\; \Delta U_1^* \, , \\
&\text{Scenario 3:} \quad \Delta U_2^* \;>\; \theta_H \;>\; \theta_L \;>\; \Delta U_1^* \, .
\end{aligned}$$

These inequalities are interpreted as follows.

Scenario 1: Individuals of any class, are better off by public good provision if their taste realization is high, $\theta_t = \theta_H$. They are worse off in case of a low taste realization, $\theta_t = \theta_L$. Scenario 1 hence gives rise to the statement that, at the *informed optimum*, *willingness to pay for the public good is independent of earning ability*, but depends only on the taste realization.

Scenario 2: As under *Scenario 1*, more productive individuals desire public good provision only if their utility gain is large, i.e. only if $\theta_2 = \theta_H$. In contrast, less productive individuals, whose utility loss is smaller, benefit from provision in any state; that is, even if $\theta_1 = \theta_L$.

Scenario 3: As under *Scenario 2*, less productive individuals always enjoy the public good. More able individuals, however, suffer from such a heavy utility loss that public good provision makes them worse off even if $\theta_2 = \theta_H$.

reversed, the conclusion of Lemma 5.1 would be reversed as well. An analysis based on this alternative premise would have to follow the same line of reasoning as the one developed below.

5.4 The twice modified Samuelson rule

In this section the optimal tax rule and the optimal provision rule for the public good – i.e. the solution to the optimization problem in Definition 5.1 – is characterized for each parameter constellation of the model; that is, for each of the three scenarios defined in the previous section. The optimal allocation rule has to satisfy incentive compatibility. Hence, I call the optimal allocation a *twice modified Samuelson rule* because it takes into account the marginal costs of public funds under a linear income tax and, in addition, the desirability of preference revelation.

5.4.1 Admissible provision rules

The optimal allocation rule is derived via a two step procedure. The *first* step solves for an optimal tax rule, taking the provision rule for the public good as given. The *second* step determines the optimal provision rule. This approach is tractable because of the fact that the *IC* constraints limit the number of admissible provision rules.

Lemma 5.2 Incentive compatible provision rules are increasing in both arguments, $\forall\theta_1$: $Q(\theta_1, \theta_L) \leq Q(\theta_1, \theta_H)$ and $\forall\theta_2 : Q(\theta_L, \theta_2) \leq Q(\theta_H, \theta_2)$.

These monotonicity constraints imply that there are only six candidate provision rules. Provision rule $Q^i : Q = 0 \Longleftrightarrow (\theta_1, \theta_2) = (\theta_L, \theta_L)$, which is part of the informed optimum, satisfies these constraints. The same is true for provision rule $Q^{i\prime}$, defined by $Q = 1 \Longleftrightarrow (\theta_1, \theta_2) = (\theta_H, \theta_H)$, provision rule Q^1, which calls for public good provision if and only if class 1 individuals have a high taste parameter $Q^1 : Q = 1 \Longleftrightarrow \theta_1 = \theta_H$, and the analogously defined provision rule $Q^2 : Q = 1 \Longleftrightarrow \theta_2 = \theta_H$. Finally, the monotonicity constraints are trivially satisfied by the constant provision rules $Q \equiv 0$ and $Q \equiv 1$.

Any such provision rule can be interpreted in terms of the influence that is assigned to individuals of different classes. For instance, under provision rule Q^i, individuals of each class have a veto against $Q = 0$: whenever at least one class of individuals collectively announces a high taste realization, then the public good is provided. Likewise, under provision rule $Q^{i\prime}$, each class has a veto against $Q = 1$. Under provision rule Q^1, the tax setting planner listens only to the preference announcement of the less able class. The more able have no influence on public good provision. Analogously, under Q^2, the decision on provision does not depend on the less able individuals' taste announcement. Finally, under $Q \equiv 0$ and $Q \equiv 1$, neither class has an impact.

5.4.2 Does incentive compatibility always matter?

As a first step, I characterize the circumstances under which the requirement of incentive compatibility indeed affects the choice of an optimal allocation rule. Obviously, incentive compatibility is not an issue if the *informed optimum*, as characterized by τ^* and Q^*, satisfies the incentive compatibility constraints in (5.3).

Under Assumption 5.1, the informed optimum has $Q^* = Q^i$. Using τ^* this implies that for class t, the informed optimum induces the following levels of indirect private utility,

$$V_t(\theta_1, \theta_2) = \begin{cases} U^*(0, w_t), & \text{if} \quad (\theta_1, \theta_2) = (\theta_L, \theta_L) \\ U^*(\tau_k, w_t) & \qquad \text{otherwise.} \end{cases}$$

These expressions can be used to check whether or not the informed optimum satisfies the constraints in (5.3). It is easily verified that IC holds if and only if, for any class t,

$$\theta_H \;\geq\; \Delta U_t^* \;\geq\; \theta_L\ .$$

This chain of inequalities is satisfied if and and only if *Scenario 1* applies. This proves the following proposition.

Proposition 5.1 The *informed optimum* is incentive compatible if and only if *Scenario 1* applies, i.e. if and only if *willingness to pay for the public good is independent of earning ability.*

Thus, there is indeed a parameter constellation where the tax setting planner gets the information on taste parameters for free; that is, without a welfare loss due to binding incentive compatibility constraints. This is the case if all high ability agents and all low ability agents want the public good to be installed only in case of a high taste realization. Put differently, whether or not an individual prefers $Q = 1$ over $Q = 0$ depends only on the taste realization but not on the ability level.

Note, however, that the absence of incentive problems is not the same as the absence of conflicting interests. To see this, suppose that the actual state of the economy is (θ_L, θ_H). In this case, the less able individuals prefer $Q = 0$, while the more productive prefer $Q = 1$. However, despite those conflicting views, neither class has an incentive to hide its true taste realization. A false announcement would not yield a preferred outcome.

Proposition 5.1 characterizes the optimal allocation rule under *Scenario 1*. Next I characterize the optimum under Scenario 2.

5.4.3 Scenario 2

The analysis proceeds in two steps. Recall that an informed utilitarian planner would choose provision rule Q^i. I first indicate what taxes a utilitarian planner chooses, given that this provision rule has to be implemented. Then, I discuss whether this provision rule remains part of an optimal allocation rule under incentive constraints.

Optimal taxes for provision rule Q^i under Scenario 2

Under Scenario 2, incentive problems arise for the following reason. Given that provision rule Q^i is chosen or implementation and taxes are such that the budget constraint binds, less able individuals will never admit a low taste realization. Hence, if the planner sticks to provision rule Q^i, a deviation from budget balance becomes unavoidable. I will now solve for the *optimal* deviation.

I first derive a concise statement of the planner's problem. From substituting Q^i into the incentive compatibility constraints in (5.3), one finds that incentive compatibility for the less able class requires that

$$\theta_H \geq V_1(\theta_L, \theta_L) - V_1(\theta_H, \theta_L) \geq \theta_L, \quad V_1(\theta_L, \theta_H) = V_1(\theta_H, \theta_H). \tag{5.5}$$

Similarly, the incentive constraints for the more productive are

$$\theta_H \geq V_2(\theta_L, \theta_L) - V_1(\theta_L, \theta_H) \geq \theta_L, \quad V_2(\theta_H, \theta_L) = V_2(\theta_H, \theta_H). \tag{5.6}$$

These incentive constraints imply that whenever $Q = 1$ the same tax rate has to be used to cover the cost of provision, $\tau(\theta_H, \theta_H) = \tau(\theta_L, \theta_H) = \tau(\theta_H, \theta_L)$.[10]

[10] As the function $U^*(\tau, w_t)$ is strictly decreasing in τ, the constraint $V_1(\theta_L, \theta_H) = V_1(\theta_H, \theta_H)$ implies $\tau(\theta_L, \theta_H) = \tau(\theta_H, \theta_H)$ and $V_2(\theta_H, \theta_L) = V_2(\theta_H, \theta_H)$ gives $\tau(\theta_H, \theta_L) = \tau(\theta_H, \theta_H)$.

This tax rate is henceforth called $\bar{\tau}$, i.e.

$$\bar{\tau} \;:=\; \tau(\theta_H, \theta_H) = \tau(\theta_L, \theta_H) = \tau(\theta_H, \theta_L) \; .$$

Analogously, define $\underline{\tau} := \tau(\theta_L, \theta_L)$. Using those tax rates, the incentive constraints for any class t can be rewritten as:

$$\theta_H \;\geq\; U^*(\underline{\tau}, w_t) - U^*(\bar{\tau}, w_t) \;\geq\; \theta_L$$

In addition, from the property of increasing differences, the private utility loss due to higher taxation is larger for individuals of class 2,

$$U^*(\underline{\tau}, w_2) - U^*(\bar{\tau}, w_2) \;>\; U^*(\underline{\tau}, w_1) - U^*(\bar{\tau}, w_1) \; .$$

Consequently, the planner only has to take the constraints $\theta_H \geq U^*(\underline{\tau}, w_2) - U^*(\bar{\tau}, w_2)$ and $U^*(\underline{\tau}, w_1) - U^*(\bar{\tau}, w_1) \geq \theta_L$ into account.

In other words, the observation that it is harder to convince the more able class that the public good should be installed implies that the *IC* conditions can be simplified. *IC* holds whenever the "poor" are willing to admit a low valuation of the public good and the "rich" are willing to admit a high valuation.

The planner's problem can now be stated in the following way: Denote by $W_p(\tau) := U^*(\tau, w_1) + U^*(\tau, w_2)$ the welfare contribution of aggregate private utility, given a tax rate of τ. An optimal choice of $\underline{\tau}$ and $\bar{\tau}$ solves the following problem, referred to as *Problem P*:

$$\begin{array}{lll}
\max_{\bar{\tau},\underline{\tau}} & p_{LL} W_p(\underline{\tau}) + (1 - p_{LL}) W_p(\bar{\tau}) & \\[2em]
\text{s.t.} & \underline{\tau} \geq 0 \; , \;\; \bar{\tau} \geq \tau_k & (\text{BC}) \; , \\[2em]
 & \theta_H \geq U^*(\underline{\tau}, w_2) - U^*(\bar{\tau}, w_2) & (\text{IC}_2) \; , \\[2em]
 & U^*(\underline{\tau}, w_1) - U^*(\bar{\tau}, w_1) \geq \theta_L & (\text{IC}_1) \; .
\end{array}$$

I denote by $\underline{\tau}^{**}$ and $\bar{\tau}^{**}$ the second best tax rates, which solve P. In addition, denote by τ_{1L} the tax rate which satisfies

$$U^*(0, w_1) - U^*(\tau_{1L}, w_1) \quad = \quad \theta_L \ .$$

τ_{1L} makes less able individuals with θ_L indifferent with respect to public good provision, given that $\underline{\tau} = 0$. Note that by the definition of scenario 1, $\tau_{1L} > \tau_k$.

Lemma 5.3

a) If $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, then $\underline{\tau}^{**} = 0$ and $\bar{\tau}^{**} = \tau_{1L}$.

b) If $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, a solution to the planner's problem exists only if there are tax rates such that IC_1 and IC_2 are binding. Furthermore, if a solution exists, then $\underline{\tau}^{**} > 0$ and $\bar{\tau}^{**} > \tau_k$.

Under *Scenario 2*, the less productive class has to be prevented from announcing a high taste parameter if in fact their taste parameter is low. In order to fix this incentive problem, the outcome $Q = 1$ is made less attractive by excessive taxation, i.e. the tax $\bar{\tau}^{**}$ exceeds the level τ_k, which would be sufficient to cover the cost of provision.

However, this excessive tax rate may generate a new incentive problem: if the public good is made less attractive, then one might end up in a situation where more productive individuals are no longer willing to admit a high valuation of the public good.

There are two possible cases. In case of a *modest incentive problem* – i.e. if $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$ – the more able still prefer $Q = 1$ under an excessive level of $\bar{\tau}$. In contrast, if incentive problems are *severe*, then Q^i is implementable only if both incentive constraints are binding. However, a pair of tax rates such that both incentive constraints are binding need not exist.[11] If those tax rates do not exist, then, under a severe incentive problem, provision rule Q^i is not implementable.

The optimal provision rule under Scenario 2

As has just been shown, if the planner wants to implement Q^i, then he has to accept the need to waste tax revenues, because the less able class loves the public good too intensively. More generally, such a waste of tax revenues becomes unavoidable if the decision on provision is made dependent on the taste announcement of the less able class. This is true for provision rules Q^i, $Q^{i\prime}$ and Q^1.

There are, however, provision rules, which do not require excessive taxes. Under provision rules $Q \equiv 0$ and $Q \equiv 1$, the decision on provision is not state dependent. Consequently, there is no need to communicate and hence no need for excessive taxes in order to ensure truth-telling. Under provision rule Q^2, the planner only has to ask only the more able class about their preferences. They want to induce public good provision if and only if $\theta_2 = \theta_H$. This implies that Q^2 can also be implemented without having to rely on excessive taxes.

A utilitarian planner faces a tradeoff. Either he sticks to provision rule Q^i and has to burn money, or he decides not to burn money but deviates from the provision rule that is part of the *informed optimum.* The following proposition shows how a utilitarian planner deals with this issue, depending on the intensity of incentive problems and his prior.

More precisely, the proposition summarizes the results from the following

[11] For standard examples of functional forms – Cobb Douglas utility in logarithmic formulation or isoelastic components of private utility – those tax rates do not exist.

exercise: For each candidate provision – i.e. for each provision rule in the set $\{Q^i, Q^{i'}, Q^1, Q^2, Q \equiv 0, Q \equiv 0\}$ – solve for the optimal tax rates that implement this provision rule under budget and incentive constraints; that is, for each candidate provision rule, solve the same kind of optimization problem as the one discussed in subsection 5.4.3 for provision rule Q^i. The solution to each optimization problem allows the computation of the resulting welfare levels, which I denote by EW^i, $EW^{i'}$, EW^1, EW^2, $EW^{Q\equiv 0}$ and $EW^{Q\equiv 1}$, respectively. A comparison of these welfare levels then determines the optimal provision rule under budget as well as incentive constraints.

The welfare maximizing provision rule depends on the prior beliefs; that is, on the probability weights that are used in the computation of expected welfare levels. I say that a provision rule can be *supported* if there exists a vector of prior beliefs p such that this provision rule turns out to be welfare maximizing.

Proposition 5.2

i) Suppose the incentive problem is *modest*, i.e. $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$. Then Q^2, Q^i and $Q \equiv 1$ can be supported.[12]

ii) Suppose the incentive problem is *severe*, $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$. Then Q^2 and $Q \equiv 1$ can be supported.

Proposition 5.2 shows that, in case of a *severe incentive problem* – i.e. if $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$ – provision rule Q^i is never chosen. Even if tax rates exist under which this rule is implementable, the planner will avoid the welfare burden of two binding incentive constraints. This shows that incen-

[12]For the special case $\theta_H = U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, provision rules Q^1 and Q^i yield the same level of expected utilitarian welfare. Otherwise Q^i is strictly superior.

tive constraints may heavily affect optimal policy: provision rule Q^i either becomes infeasible or undesirable.

In case of a *severe incentive problem*, only provision rules Q^2 and $Q \equiv 1$, which avoid excessive taxation by not giving any influence to the less able class, are candidates for maximizing expected utilitarian welfare.[13] In contrast, under *modest incentive problems*, the planner may stick to provision rule Q^i, i.e. the burden of excessive taxation is not necessarily prohibitive.[14]

5.4.4 Scenario 3

Under scenario 3, the incentive problem known from scenario 2 – i.e. that under budget balance, the less able will never admit a low valuation because they want to ensure $Q = 1$ – is accompanied by another incentive problem that is due to the more able class of individuals. Under budget balance, they will not admit a high valuation because they hate having to pay for the public good. Consequently, a case of modest incentive problems is not possible under scenario 3.

To see this, recall that the incentive problem caused by the less able requires that the public good becomes less attractive. This requires an excessive tax rate whenever the public good is installed. By contrast, the incentive problem caused by the more able class calls for an excessive tax rate that makes

[13] Which of them is superior depends on the likelihood of the states in which they implement a different allocation. By definition of scenario 1, Q^2 yields ex post the higher welfare level if $(\theta_1, \theta_2) = (\theta_L, \theta_L)$, rule $Q \equiv 1$ is superior if $(\theta_1, \theta_2) = (\theta_H, \theta_L)$.

[14] To see why Q^i comes in as an additional candidate for optimal policy, suppose that p_{LL} is sufficiently large in the sense that Q^2 yields a higher expected welfare level than $Q \equiv 1$. Ex post, rule Q^i gives a higher welfare level as compared to rule Q^2 if $(\theta_1, \theta_2) = (\theta_H, \theta_L)$, and rule Q^2 is more attractive if $\theta_2 = \theta_H$. Hence, to ensure optimality of rule Q^i conditionally on $(\theta_1, \theta_2) \neq (\theta_L, \theta_L)$, the probability that $\theta_2 = \theta_H$ must be small.

non-provision a less attractive outcome. Those two incentive problems aggravate each other: Trying to fix the incentive problem for the less able class makes the outcome $Q = 1$ less attractive, and this implies that it becomes even harder to make the more able class willing to admit a high taste realization.

Consequently, whenever a provision rule is chosen that gives influence to both classes – recall that this is the case under Q^i and $Q^{i'}$– one ends up with two binding incentive constraints. As has been discussed in the previous subsection this may imply that these provision rules can no longer be implemented. This is the case if no pair of tax rates exists that makes the incentive constraints of both classes binding.

One easily verifies, however, that even if implementation is possible, those provision rules become undesirable. Put differently, there do not exist prior beliefs that support Q^i or $Q^{i'}$ under Scenario 3. The reason has already been discussed in the previous subsection. The welfare burden of two excessive tax rates becomes prohibitive. For instance, one can show that the planner prefers to communicate with only one class of individuals; that is, to choose Q^1 or Q^2 instead. Even though this implies that some information is lost, it requires only one excessive tax rate. The total effect is a larger level of expected utilitarian welfare.

5.5 Robustness

Up to now, it has been assumed that there are only two alternative uses of the proceeds from linear taxation: covering the cost of public good provision; and "waste" for incentive reasons. This raises the question how far the results depend on these assumptions: e.g. what would the the analysis look like if direct income transfers were allowed and excessive tax revenues could be returned to the agents? A further concern is the degree to which results rely on the fact that taxation is linear. Below, I report on a series of exercises that study the robustness of this paper's results.

I first argue that the linear tax on income is necessary if a higher tax revenue requirement is to do greater harm to the more productive. To see this, suppose that taxation is non-linear and that direct income transfers are allowed as in Chapter 4.[15] In this setting, an increased revenue requirement, due to public good provision, yields a larger private utility loss for less able individuals. The property of increasing differences, established in Lemma 5.1, is hence replaced by *decreasing differences*. The underlying reason is as follows: An optimal non-linear income tax is an arrangement of redistribution under incentive constraints. As shown in Chapter 4, this implies that the more able class who already finances the transfer system cannot be used to generate additional tax revenues for the public good. In addition, it is proven that, even though incentive requirements imply that a utilitarian planner has to deviate from the *complete information benchmark*, the more flexible instrument of a non-linear income tax make it possible to avoid a waste of tax revenues.

Similar results can be obtained in the following environment: the linear income tax rate τ is not only raised to cover the cost of public good provision, but also in order to finance a lump sum transfer α, which is equal for all individuals in the economy. If both τ and α are set optimally according to a utilitarian welfare function,[16] the need to finance a public project has two opposing effects. It leads to an increase of τ and to a reduction of α. It is

[15] The problem of preference revelation is introduced into the framework of a two-class economy, as, for instance, analyzed in Stiglitz (1982) or Boadway and Keen (1993).

[16] This model has been studied in more detail by Sheshinski (1972) and Hellwig (1986).

easy to find examples where the second effect dominates,[17] implying, once again, that one has decreasing differences. Hence, the mere fact that a linear tax system is in operation does not imply that the rich are going to suffer relatively more if additional tax revenues are needed. This effect will occur only if redistribution is not reduced too much in response to public good provision.

To sum up, the discussion shows that the analysis of this paper is applicable only if the level of income transfers is kept fixed and public good provision is financed by *additional* linear taxes.

5.6 Concluding Remarks

An important theme in political economics is the question whether or not political competition induces efficient outcomes.[18] Polo (1998) and Svensson (2000) study this question in the context of a probabilistic voting model. The authors show that equilibrium rents, defined as an excess of tax revenues over the cost of public good provision, are positive. This observation seems to support the view that political competition yields undesirable outcomes. The present paper, by contrast, does not attempt to model the outcome of the political process in a somewhat realistic manner. Instead, the paper describes what an ideal arrangement looks like under the assumptions that there exist two groups of agents and that a policy decision has to based on

[17] Suppose that $u(x) = \ln(x)$ and $v(x) = x^{\gamma}$. For $\gamma \to 1$, the optimal linear tax rate becomes independent of the revenue requirement $kQ(\theta_1, \theta_2)$; that is, the public good is just crowding out the income transfer α.

[18] Alternative views on that issue are associated with the labels of *Virginia*, for emphasis on inefficiencies induced by rent-seeking, and *Chicago*, for the idea that competition among politicians leaves no room for rent-extraction. See Coate and Morris (1995) and Persson and Tabellini (2000) for further discussion.

their behavior in a revelation game. It turns out that a constrained efficient allocation may involve a waste of tax revenues. Burning money may be a valuable policy option in order to ensure that private information becomes available for public decision making.

This sheds a different light on the results mentioned above. The mere fact that an equilibrium outcome of some political game involves a budget surplus cannot be taken as evidence of inefficiency. Such a conclusion requires the identification of the constrained efficient allocations under all relevant informational, institutional and technological restrictions. In particular, the present paper suggests that excessive tax revenues may not result from the deficiencies political competition but from the fact that the set of policy instruments is rather limited.

5.7 Appendix

Proof of Lemma 5.1: Straightforward calculations yield

$$Y_t'(\tau) = \frac{u'((1-\tau)Y_t(\tau)) + (1-\tau)Y_t(\tau)u''((1-\tau)Y_t(\tau))}{(1-\tau)^2 u''((1-\tau)Y_t(\tau)) - w_t^{-2} v''\left(Y_t(\tau)/w_t\right)}\ , \quad \text{and}$$

$$\frac{\partial^2 U^*(\tau, w_t)}{\partial\tau\partial w_t} = \frac{Y_t'(\tau)}{w_t^2}\left[v'\left(\frac{Y_t(\tau)}{w_t}\right) + \frac{Y_t(\tau)}{w_t}\, v''\left(\frac{Y_t(\tau)}{w_t}\right)\right].$$

If this cross derivative is negative, then preferences satisfy the property of increasing differences.

■

Proof of Lemma 5.2: The monotonicity of admissible provision rules is derived as follows: consider for example the two incentive compatibility constraints for class 1, given that $\theta_2 = \theta_L$:

$$\text{if } \theta_1 = \theta_L: \quad \theta_L Q(\theta_L, \theta_L) + V_1(\theta_L, \theta_L) \geq \theta_L Q(\theta_H, \theta_L) + V_1(\theta_H, \theta_L) \; ,$$

$$\text{if } \theta_1 = \theta_H: \quad \theta_H Q(\theta_H, \theta_L) + V_1(\theta_H, \theta_L) \geq \theta_H Q(\theta_L, \theta_L) + V_1(\theta_L, \theta_L) \; .$$

Adding up these inequalities gives $Q(\theta_H, \theta_L) \geq Q(\theta_L, \theta_L)$. Similarly, one derives the constraints $Q(\theta_L, \theta_H) \geq Q(\theta_L, \theta_L)$, $Q(\theta_H, \theta_H) \geq Q(\theta_H, \theta_L)$ and $Q(\theta_H, \theta_H) \geq Q(\theta_L, \theta_H)$.

■

Proof of Lemma 5.3:

a) Suppose $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$. Consider the relaxed problem, which ignores IC_2. At a solution, it has to be true that $\underline{\tau}^{**} = 0$. Otherwise $\underline{\tau}^{**}$ could be reduced in a feasible and incentive compatible manner, thereby contradicting optimality. The optimal level of $\bar{\tau}$ then has to ensure that IC_1 is binding. Hence, $\bar{\tau}^{**} = \tau_{1L}$. The assumption $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$ implies that IC_2 can indeed be ignored.

b) Let $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, and suppose a solution exists. Let τ_{2H} be the tax rate that is implicitly defined by the equation $\theta_H = U^*(0, w_2) - U^*(\tau_{2H}, w_2)$. Again, by definition of Scenario 1, $\tau_{2H} > \tau_k$. Note that $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$ implies $\tau_{2H} < \tau_{1L}$.

   i) It must be the case that $\bar{\tau}^{**} > \tau_k$. Suppose, to the contrary, that $\bar{\tau}^{**} = \tau_k$. Then IC_1 is violated. To see this, note that $\bar{\tau}^{**} = \tau_k$ implies

$$U^*(\underline{\tau}^{**}, w_1) - U^*(\bar{\tau}^{**}, w_1) \quad \leq \quad \Delta U_1^* \quad < \quad \theta_L \; .$$

ii) It must be the case that $\underline{\tau}^{**} > 0$. Suppose, to the contrary, that $\underline{\tau}^{**} = 0$. Then IC_1 implies $\bar{\tau}^{**} \geq \tau_{1L}$, and IC_2 implies $\bar{\tau}^{**} \leq \tau_{2H}$, contradicting $\tau_{2H} < \tau_{1L}$.

iii) At least one (IC) constraint has to be binding. Otherwise – with $\underline{\tau}^{**} > 0$ and $\bar{\tau}^{**} > \tau_k$ – both tax rates could be reduced in a feasible and incentive compatible manner. To see that both (IC) constraints have to be binding, suppose, for instance, that, at an optimum, IC_2 binds and IC_1 does not. Then, both tax rates could be reduced in a feasible and incentive compatible manner – keeping the equality in the constraint for class 2, while not violating the one for class 1 – thereby increasing utilitarian welfare.

■

Proof of Proposition 5.2: First, optimal welfare EW^q for each provision rule is derived, where the superscript q refers to the provision rule. In the second step, those welfare levels are compared to determine the optimal provision rule.

Rule $Q \equiv 0$: $EW^{Q\equiv 0} = U^*(0, w_1) + U^*(0, w_2)$.

Rule $Q^{i\prime}$: Along the same lines as for Problem P, one derives that the planner has to solve the following problem if $Q^{i\prime}$ is chosen: choose $\bar{\tau}$ and $\underline{\tau}$ n order to maximize $(1 - p_{HH})W_p(\underline{\tau}) + p_{HH}W_p(\bar{\tau})$ subject to the following set of constraints, the budget constraints $((BC))$ $\underline{\tau} \geq 0$ and $\bar{\tau} \geq \tau_k$ and the incentive constraints

$$\theta_H \geq U^*(\underline{\tau}, w_2) - U^*(\bar{\tau}, w_2) \quad (\text{IC}_2) \ ,$$

and

$$U^*(\underline{\tau}, w_1) - U^*(\bar{\tau}, w_1) \geq \theta_L \quad (\mathrm{IC}_1) \ .$$

The solution to this problem has been characterized in Lemma 5.3. If $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, then expected welfare is given as:

$$\begin{aligned} EW^{i\prime} &= (1 - p_{HH})(U^*(0, w_1) + U^*(0, w_2)) \\ &\quad + p_{HH}(2\theta_H + U^*(\tau_{1L}, w_1) + U^*(\tau_{1L}, w_2)) \ . \end{aligned} \tag{5.7}$$

If $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, either rule $Q^{i\prime}$ cannot be implemented or the optimal combination of $\underline{\tau}$ and $\bar{\tau}$, for which both IC constraints are binding, is chosen. Denote these as τ_{LL} and τ_{HH}, respectively. Then,

$$\begin{aligned} EW^{i\prime} &= (1 - p_{HH})(U^*(\tau_{LL}, w_1) + U^*(\tau_{LL}, w_2)) \\ &\quad + p_{HH}(2\theta_H + U^*(\tau_{HH}, w_1) + U^*(\tau_{HH}, w_2)) \ . \end{aligned} \tag{5.8}$$

Rule Q^1: Under provision rule Q^1, (IC_2) requires $V_2(\theta_L, \theta_H) = V_2(\theta_L, \theta_L)$ and $V_2(\theta_H, \theta_H) = V_2(\theta_H, \theta_L)$, or equivalently $\underline{\tau} := \tau(\theta_L, \theta_L) = \tau(\theta_L, \theta_H)$ and $\bar{\tau} := \tau(\theta_H, \theta_H) = \tau(\theta_H, \theta_L)$. The planner's problem becomes:

$$\max_{\bar{\tau}, \underline{\tau}} \quad (p_{LL} + p_{LH}) W_p(\underline{\tau}) + (p_{HL} + p_{HH}) W_p(\bar{\tau})$$

$$\text{s.t.} \qquad \underline{\tau} \geq 0 \ , \quad \bar{\tau} \geq \tau_k \qquad\qquad (\mathrm{BC}) \ ,$$

$$\theta_H \geq U^*(\underline{\tau}, w_1) - U^*(\bar{\tau}, w_1) \geq \theta_L \qquad\qquad (\mathrm{IC}_1) \ .$$

It is easily verified that, at an optimum, only the constraint $U^*(\underline{\tau}, w_1) - U^*(\bar{\tau}, w_1) \geq \theta_L$ is binding. Optimal taxes are given as $\underline{\tau}^{**} = 0$ and $\bar{\tau}^{**} = \tau_{1L}$.

Expected utilitarian welfare under rule Q^1 equals:

$$\begin{aligned} EW^1 &= (p_{LL} + p_{LH})(U^*(0, w_1) + U^*(0, w_2)) \\ & (p_{HL} + p_{HH})(U^*(\tau_{1L}, w_1) + U^*(\tau_{1L}, w_2)) \\ & + p_{HL}(\theta_L + \theta_H) + 2p_{HH}\theta_H \ . \end{aligned} \tag{5.9}$$

Rule Q^2: Under provision rule Q^2, (IC_1) requires $V_1(\theta_H, \theta_L) = V_1(\theta_L, \theta_L)$ and $V_1(\theta_H, \theta_H) = V_1(\theta_L, \theta_H)$, or equivalently $\underline{\tau} := \tau(\theta_L, \theta_L) = \tau(\theta_H, \theta_L)$ and $\bar{\tau} := \tau(\theta_H, \theta_H) = \tau(\theta_L, \theta_H)$. The planner's problem can be written as:

$$\begin{aligned} \max_{\bar{\tau}, \underline{\tau}} \quad & (p_{LL} + p_{HL})W_p(\underline{\tau}) + (p_{LH} + p_{HH})W_p(\bar{\tau}) \\ \text{s.t.} \quad & \underline{\tau} \geq 0 \ , \quad \bar{\tau} \geq \tau_k \qquad \text{(BC)} \ , \\ & \theta_H \geq U^*(\underline{\tau}, w_2) - U^*(\bar{\tau}, w_2) \geq \theta_L \qquad (\text{IC}_2) \ . \end{aligned}$$

By definition of scenario 2, the optimal tax policy $\tau = 0$ if $Q = 0$ and $\tau = \tau_k$ if $Q = 1$ is incentive compatible, and expected utilitarian welfare becomes:

$$\begin{aligned} EW^2 &= (p_{LL} + p_{HL})(U^*(0, w_1) + U^*(0, w_2)) \\ & + (p_{LH} + p_{HH})(U^*(\tau_k, w_1) + U^*(\tau_k, w_2)) \\ & + p_{LH}(\theta_L + \theta_H) + 2p_{HH}\theta_H \ . \end{aligned} \tag{5.10}$$

Rule Q^i: The solution to this problem has been characterized in Lemma 5.3, i.e. if $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L})$, expected utilitarian equals:

$$\begin{aligned} EW^i &= p_{LL}(U^*(0,w_1) + U^*(0,w_2)) \\ &+(1-p_{LL})(U^*(\tau_{1L},w_1) + U^*(\tau_{1L},w_2)) \\ &+(p_{LH}+p_{HL})(\theta_L+\theta_H) + 2p_{HH}\theta_H \ . \end{aligned} \tag{5.11}$$

If, to the contrary, $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, then, if a solution to the planner's problem exists, expected utilitarian welfare equals:

$$\begin{aligned} EW^i &= p_{LL}(U^*(\tau_{LL},w_1) + U^*(\tau_{LL},w_2)) \\ &+(1-p_{LL})(U^*(\tau_{HH},w_1) + U^*(\tau_{HH},w_2)) \\ &+(p_{LH}+p_{HL})(\theta_L+\theta_H) + 2p_{HH}\theta_H \ . \end{aligned} \tag{5.12}$$

Rule $Q \equiv 1$: Under provision rule $Q \equiv 1$, expected utilitarian welfare equals:

$$\begin{aligned} EW^{Q\equiv 1} &= U^*(\tau_k,w_1) + U^*(\tau_k,w_2) \\ &+2p_{LL}\theta_L + (p_{LH}+p_{HL})(\theta_L+\theta_H) + 2p_{HH}\theta_H \ . \end{aligned} \tag{5.13}$$

The proof of Proposition 5.2 is now established by the following claims:

Claim 1. Under scenario 2, $Q \equiv 0$ and $Q^{i\prime}$ are strictly dominated by Q^2.

$$\begin{aligned} EW^2 &> (1-p_{HH})(U^*(0,w_1) + U^*(0,w_2)) \\ &+p_{HH}(U^*(\tau_k,w_1) + U^*(\tau_k,w_2) + 2\theta_H) \ . \end{aligned} \tag{5.14}$$

Under scenario 2, the right hand side is strictly larger than $EW^{i'}$ and $EW^{Q\equiv 0}$.

Claim 2. If $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, then $EW^{i'} = EW^i$: Using the equations $\theta_H = U^*(\tau_{LL}, w_2) - U^*(\tau_{HH}, w_2)$ and $U^*(\tau_{LL}, w_1) - U^*(\tau_{HH}, w_1) = \theta_L$ to substitute for θ_L and θ_H in the expressions for $EW^{i'}$ and EW^i in equations (5.12) and (5.8) reveals that $EW^{i'} = EW^i$.

Claim 3. If $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, then $EW^i < EW^2$: This is a direct consequence of Claims 1 and 2.

Claim 4. If $\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, then $EW^i \geq EW^1$ with equality if and only if $\theta_H = U^*(0, w_2) - U^*(\tau_{1L}, w_2)$: To see this, use equations (5.11) and (5.9), as well as the definition of τ_{1L}, to derive:

$$EW^i - EW^1 = p_{LH}(U^*(\tau_{1L}, w_2) + \theta_H - U^*(0, w_2)) \,. \tag{5.15}$$

Claim 5. If $\theta_H < U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, then $EW^1 < EW^2$: Equations (5.9) and (5.10) imply:

$$\begin{aligned}
&EW^2 - EW^1 = \\
&p_{HL}(U^*(0, w_2) - U^*(\tau_{1L}, w_2) - \theta_H) \\
&+p_{LH}(U^*(\tau_k, w_1) + U^*(\tau_k, w_2) + \theta_L + \theta_H - U^*(0, w_1) - U^*(0, w_2)) \\
&+p_{HH}(U^*(\tau_k, w_1) + U^*(\tau_k, w_2) - U^*(\tau_{1L}, w_1) - U^*(\tau_{1L}, w_2)) \,.
\end{aligned}$$

All terms in this sum are strictly positive under Scenario 2.

Claim 6. $EW^2 - EW^{Q\equiv 1}$ may become positive or negative, depending on the prior probabilities: From equations (5.13) and (5.10):

$$EW^2 - EW^{Q\equiv 1} =$$

$$p_{LL}(U^*(0,w_1) + U^*(0,w_2) - U^*(\tau_k,w_1) - U^*(\tau_k,w_2) - 2\theta_L)$$

$$+p_{HL}(U^*(0,w_1) + U^*(0,w_2) - U^*(\tau_k,w_1) - U^*(\tau_k,w_2) - \theta_L - \theta_H)\ .$$

Under scenario 2, the first term is positive and the second is negative.

Claim 7. Let $\theta_H \geq U^*(0,w_2) - U^*(\tau_{1L},w_2)$. $EW^2 - EW^i$ and $EW^{Q\equiv 1} - EW^i$ may become positive or negative, depending on the prior probabilities: From equations (5.10), (5.11) and (5.13), one derives:

$$EW^2 - EW^i =$$

$$p_{HL}(U^*(0,w_2) - U^*(\tau_{1L},w_2) - \theta_H)$$

$$+(p_{LH} + p_{HH})(U^*(\tau_k,w_1) + U^*(\tau_k,w_2) - U^*(\tau_{1L},w_1) - U^*(\tau_{1L},w_2))\ ,$$

and

$$EW^{Q\equiv 1} - EW^i =$$

$$p_{LL}(U^*(\tau_k,w_1) + U^*(\tau_k,w_2) + 2\theta_L - U^*(0,w_1) - U^*(0,w_2))$$

$$+(1 - p_{LL})(U^*(\tau_k,w_1) + U^*(\tau_k,w_2) - U^*(\tau_{1L},w_1) - U^*(\tau_{1L},w_2))\ .$$

For both differences the first term is negative and the second is positive. The expressions for $EW^2 - EW^{Q\equiv 1}$, $EW^2 - EW^i$ and $EW^{Q\equiv 1} - EW^i$ derived above are linear, hence continuous in the probabilities, implying that, for

$\theta_H \geq U^*(0, w_2) - U^*(\tau_{1L}, w_2)$, any of the rules Q^2, Q^i or $Q \equiv 1$ may yield the maximal level of welfare.

$\blacksquare$

Bibliography

Al-Najjar, N. (2004). Aggregation and the law of large numbers in large economies. *Games and Economic Behavior*, 47:1–35.

Alòs-Ferrer, C. (2002). Individual randomness in economic models with a continuum of agents. Working Paper 9807, Dpt. of Economics, University of Vienna.

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

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.

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

Auriol, E. and Gary-Bobo, R. (2005). On robust constitution design. Mimeo, Universit Paris 1 Panthon-Sorbonne.

Austen-Smith, D. and Banks, J. (1996). Information aggregation, rationality and the Condorcet jury theorem. *American Political Science Review*, 90:34–45.

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

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

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

Bierbrauer, F. (2005a). Collectively incentive compatible tax systems. Mimeo, Max Planck Institute for Research on Collective Goods, Bonn.

Bierbrauer, F. (2005b). Distortionary taxation and the free-rider problem. Mimeo, Max Planck Institute for Research on Collective Goods, Bonn.

Bierbrauer, F. (2005c). Optimal income taxation and public good provision in a two-class economy. Preprint 2005/25, Max Planck Institute for Research on Collective Goods, Bonn.

Bierbrauer, F. and Sahm, M. (2005). Informative voting and the Samuelson rule. Mimeo, Max Planck Institute for Research on Collective Goods, Bonn and LMU, München.

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.

Boylan, T. (1998). Coalition-proof implementation. *Journal of Economic Theory*, 82:132–143.

Chung, K. and Ely, J. (2003). Ex-post incentive compatible mechanism design. Mimeo, Northwestern University.

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

Clarke, E. (1971). Multipart pricing of public goods. *Public Choice*, 11:17–33.

Coate, S. and Morris, S. (1995). On the form of transfers to special interests. *Journal of Political Economy*, 103:1210–1235.

Crawford, V. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50:1431–1451.

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.

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

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

Dierker, E. and Haller, H. (1990). Tax systems and direct mechanisms in large finite economies. *Journal of Economics*, 52:99–116.

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

Feddersen, T. and Pesendorfer, W. (1997). Voting behaviour and information aggregation in elections with private information. *Econometrica*, 65:1029–1058.

Feldman, M. and Gilles, C. (1985). An expository note on individual risk without aggregate uncertainty. *Journal of Economic Theory*, 35:26–32.

Gary-Bobo, R. and Jaaidane, V. (2000). Polling mechanisms and the demand revelation problem. *Journal of Public Economics*, 76:203–238.

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.

Gersbach, H. (2005). Democratic mechanisms: Double majority rules and flexible agenda costs. Mimeo, University of Heidelberg.

Green, J. and Laffont, J. (1977). Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica*, 45:472–487.

Green, J. and Laffont, J. (1979). *Incentives in Public Decision-Making.* North-Holland Publishing Company.

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

Groves, T. (1973). Incentives in teams. *Econometrica*, 41:617–663.

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

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

Hamilton, J. and Slutsky, S. (2005). Optimal nonlinear income taxation with a finite population. *Journal of Economic Theory*, forthcoming.

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

Hellwig, M. (1986). The optimal linear income tax revisited. *Journal of Public Economics*, 31:163–179.

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

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

Hellwig, M. (2005a). Optimal utilitarian income taxation with finitely many types. Mimeo, Max Planck Institute for Research on Collective Goods, Bonn.

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

Heuser, H. (1998). *Lehrbuch der Analysis, Teil 1.* Teubner; Stuttgart, Leipzig.

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. (1987). Incentives and the allocation of public goods. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*. North-Holland, Amsterdam.

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.

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.

Meirowitz, A. (2005). Polling games and information revelation in the downsian framework. *Games and Economic Behavior*, 51:464–489.

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

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

Moore, J. (1992). Implementation, contracts, and renegotiation in environments with complete information. In Laffont, J.-J., editor, *Advances in Economic Theory: Sixth World Congress, vol. I*. Cambridge, UK, Cambridge University Press.

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

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.

Norman, P. (2004). Efficient mechanisms for public goods with use exclusion. *Review of Economic Studies*, 71:1163–1188.

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

Piketty, T. (1993). Implementation of first-best allocations via generalized tax schedules. *Journal of Economic Theory*, 61:23–41.

Piketty, T. (1999). The information-aggregation approach to political institutions. *European Economic Review*, 43:791–800.

Polo, M. (1998). Electoral competition and political rents. Mimeo, Innocenzo Gasparini Institute for Economic Research, Milan.

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

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

Schmitz, P. (1997). Monopolistic provision of excludable public goods under private information. *Public Finance/ Finance Publiques*, 52:89–101.

Schultz, C. (1996). Polarization and inefficient policies. *Review of Economic Studies*, 63:331–343.

Sheshinski, E. (1972). The optimal linear income-tax. *Review of Economic Studies*, 39:297–302.

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

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

Svensson, J. (2000). Controlling spending: Electoral competition, polarization and endogenous platforms. Mimeo, Institute for International Economic Studies, Stockholm University.

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

Weymark, J. (1987). Comparative static properties of optimal nonlinear income taxes. *Econometrica*, 55:1165–1185.

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

Erklärung

Hiermit erkläre ich, dass ich die Dissertation selbständig angefertigt und mich anderer als der in ihr angebenen Hilfsmittel nicht bedient habe, insbesondere, dass aus anderen Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche gekennzeichnet und mit Quellenangaben versehen sind, nicht stattgefunden haben.

Felix Bierbrauer

Curriculum Vitae

since 2004	Research assistant, *Max Planck Institute for Research on Collective Goods*, Bonn
2002 – 2004	Research and teaching assistant at the chair for economic theory, Professor Martin Hellwig, Ph.D., *University of Mannheim*
since 2002	Doctoral student, *Center for Doctoral Studies in Economics and Management, University of Mannheim*
09/2002	Diploma in Economics, *University of Mannheim*
1997 – 2002	Studies in economics at the universities of Saarbrücken, Mannheim and Tilburg, The Netherlands
1996-1997	Civilian national service ("Zivildienst"), Jugendzentrum Försterstraße, Saarbrücken
1996	Graduation from secondary school ("Abitur"), Gymnasium am Krebsberg, Neunkirchen/ Saar
1977	Born in Püttlingen