Optimal Democratic Mechanisms for Taxation and Public Good Provision*

Felix Bierbrauer† Marco Sahm**
MPI, Bonn TU, München

This version: November 21, 2008

Abstract

We study the interdependence of optimal tax and expenditure policies. An optimal policy requires that information on preferences is made available. We first study this problem from a general mechanism design perspective and show that efficiency is possible only if the individuals who decide on public good provision face an own incentive scheme that differs from the tax system. We then study democratic mechanisms with the property that tax payers vote over public goods. Under such a mechanism, efficiency cannot be reached and welfare from public good provision declines as the inequality between rich and poor individuals increases.

Keywords: Public goods, optimal taxation, two-dimensional heterogeneity, asymmetric information

JEL: H41, D71, D72, D82

*We are very grateful for comments by Thomas Gaube, Mike Golosov, Aleh Tsyvinski, Martin Hellwig, and Klaus Schmidt and John Weymark. We also thank seminar participants at the Massachusetts Institute of Technology, Cambridge, MA, the Ludwig-Maximilians-Universität, München, the Max Planck Institute for Research on Collective Goods, Bonn and the Humboldt University, Berlin. The first author also thanks the Massachusetts Institute of Technology for its hospitality.

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

**Technical University of Munich, Lehrstuhl für Volkswirtschaftslehre, Arcisstraße 21, 80333 München, Germany. Email: marco.sahm@wi.tum.de

1 Introduction

This paper studies the interdependence of optimal tax policies and optimal expenditure policies, with a focus on information and incentive problems. An interdependence arises under the assumption that tax revenues are used to pay for a public good. An individual's assessment of a public good will then depend not only on the individual's preferences for public good but also on his earning abilities which determine his income. Hence, the tax policy affects an individual's willingness to communicate his public goods preferences truthfully to "the system". This raises the question what an optimal tax and expenditure policy looks like that takes this interdependence into account.

We address this question from two different angles. First, we study optimal tax and expenditure policies from a mechanism design perspective. In this approach, the only constraints that an allocation has to satisfy are physical feasibility and incentive compatibility, i.e., it has to be taken into account that individuals are privately informed about their abilities and their preferences. Second, we study a class of allocation mechanisms that we call *democratic* mechanisms. These mechanisms take, in addition, two institutional constraints into account: (i) individuals vote over public goods and (ii) all individuals pay taxes.

Under a general allocation mechanism individuals communicate their utility functions. Decisions on public good provision can thus reflect the preference intensities of individuals. By contrast, under a voting procedure individuals can only express whether they want a public good to be provided. If so, they vote *yes* and otherwise they vote *no*. Hence, under a voting scheme, each person has the same influence on public good provision, irrespective of the person's utility function.

A general allocation mechanism makes it possible to separate the individuals who pay taxes from the individuals who decide on public good provision. For instance, it is possible to draw a random sample of individuals who are exempt from the general tax system and whose public goods preferences are elicited via some Clarke-Groves mechanism. The decision on public good provision would then be a function of the characteristics of those individuals who are in the sample. A democratic mechanism excludes this possibility. It is based on the normative premise that the relevant preferences for a decision on public good provision are the preferences of those who pay for the public good via the tax system. We refer to this principle as "no taxation without representation."

The assumptions that individuals express their preferences via voting and that all individuals are taxpayers are made in many papers on the political economy of taxation and public good provision. They are empirically motivated. In a democratic society, citizens express political support by means of voting decisions. Moreover, a constitutional principle in a democracy is that citizens with the same characteristics are treated equally. Suppose there are two households who have the same characteristics. In a democracy, it would not be acceptable that one of them decides on public good

provision whereas the other household works in order to generate the tax revenues that are needed to pay for the public good. Rather, both households should have the same influence on public good provision and the same duty to contribute to its financing.
We introduce these assumptions into an analysis of optimal tax and expenditure policies. The key difference to a positive political economy analysis is that neither the tax system nor the rule according to which votes are translated into expenditures are determined as the equilibrium of a political game. Instead, the tax system and the provision rule for the public good are chosen in order to maximize welfare. Our analysis thus determines the optimal allocation that satisfies a specific political economy constraint, namely, that the preferences of taxpayers are decisive for public good provision and that these preferences are articulated via a voting system.

We consider a model with a continuum of individuals who differ in their abilities and either have a high or a low preference for the public good. The public good is financed via an optimal nonlinear income tax. An information problem arises because there is uncertainty about the population share of individuals with a high taste parameter. For this environment, we establish the following results:
First, optimal mechanism design makes it possible to achieve an efficient provision of public goods; i.e., the incentive constraints due to private information on public goods preferences do not affect the optimal allocation. To establish this result, we consider a finite random sample of individuals. These individuals face an own incentive scheme that is different from the optimal income tax and ensures that truth-telling of individuals in the sample is a dominant strategy. If we let the sample size go to infinity, the sample provides very accurate information about the distribution of public goods preferences in the economy as a whole and the decision on public good provision is approximately efficient.
Second, with a democratic mechanism an efficient provision of public goods can not be approximated. This result is based on the observation that, under an optimal income tax system, an individual's valuation of the public good is an increasing function of the individual's taste for public goods and of the individual's skill level. Highly productive individuals have, ceteris paribus, a larger valuation because they do not suffer as much if taxes are raised to pay for the public good. This dependence of individual valuations on abilities is the driving force behind the failure of efficiency: If many individuals in the economy have a low preference for the public good, this implies that every high-skilled individual – even those with a low taste for public goods – has an above average valuation of the public good. Consequently, all high-skilled individuals vote in favor of more public spending. From a welfare perspective, this generates an excessive demand for public goods because a comparison of social costs and benefits would take the high tax burden of low-skilled individuals into account. Analogously, if many individuals have a high valuation of the public good, then all low-skilled individuals have a below-average valuation and will thus vote for less public spending. The demand for public goods is then too low because the lower tax burden of high skilled individuals is not

given appropriate weight. The consequence of these considerations is that the provision rule for the public good has to be distorted in order to eliminate both the excessive demand for public goods of high-skilled individuals and the deficient demand of low skilled individuals.

Finally, we show that the discrepancy between an efficient provision rule and the optimal democratic mechanism is an increasing function of the extent of skill heterogeneity: the distortions of the optimal democratic mechanism become larger as the difference between a high-skilled individual and a low-skilled individual increases. This is our main insight: Under a democratic mechanism, inequality is bad not only for distributive reasons, but also because it is harmful for public good provision.

Our paper also makes a technical contribution. The attempt to link the theory of optimal taxation with the theory of public goods faces a conceptual problem. The theory of optimal income taxation, in the tradition of Mirrlees (1971), analyzes "large" economies in which every individual takes the tax system and expenditures as given. The theory of public goods, in the tradition of Clarke (1971) and Groves (1973), by contrast, is based on a "small" economy in which every individual can affect how much of a public good is provided.

This difference between these environments confronts us with a modeling choice. We could either introduce skill heterogeneity and distributive concerns into a finite economy model, or, we could introduce heterogeneity in public goods preferences and uncertainty about the social benefits of public goods provision into a large economy model of the Mirrleesian type. We have chosen the latter route, for the following reasons. First, we are interested in analyzing a situation where the general tax system is used to cover the cost of public goods provision. Empirically, this assumption seems plausible for public goods that are provided at a national scale such as national defense or the judicial system and which are consumed jointly by millions of individuals. For such a setting, an analysis which assumes that individuals chose their labor supply without taking repercussions for aggregate tax revenues and public expenditures into account seems like a reasonable approximation. Also, for the large economy model, we are able to characterize an optimal mechanism for optimal taxation and public goods provision, which can serve as a benchmark outcome for the study of democratic mechanisms.[1]

To deal with the problem of preference elicitation in a large economy, we adopt an idea that has been developed by Green and Laffont (1979). We consider a continuum of agents but assume that a finite random sample of individuals is drawn for the purpose of preference elicitation. Under a general mechanism, these individuals interact according to some revelation game, whereas under a democratic mechanism, we consider a random sample of tax payers who vote over the public good. To single out the

[1] By contrast, the attempt to characterize an optimal mechanism for a finite economy model in which individuals differ in preferences and abilities leads to a more difficult multi-dimensional mechanism design problem whose solution would be an important contribution in itself. The solution of this problem is beyond the scope of this paper.

"reasonable" equilibrium in an economy with a continuum of individuals, we study the properties of these mechanisms as the sample size goes to infinity.

Our paper is related to various strands of the literature.

It contributes to a recent literature that uses mechanism design methods to characterize optimal tax policies. See, for instance, Golosov et al. (2003), or Kocherlakota (2005). A paper that is related to our work is Acemoglu et al. (2007), who also introduce a political economy constraint into a model of optimal taxation. In particular, they relax the assumption of a benevolent mechanism designer and assume that citizens can use a voting mechanism to control politicians.

Bassetto and Phelan (2008) develop a model that is similar to ours, even though their paper differs in focus. They characterize an optimal insurance mechanism for an exchange economy with aggregate uncertainty. The authors focus on the characterization of an optimal insurance mechanism that is robust in the sense that it does not require strong assumptions on individual beliefs nor suffers from multiple equilibria. Our work is related in that we model aggregate uncertainty in a similar way, and also address the problem of robustness. While we use a a sampling approach for the purpose of equilibrium selection, we also focus on implementation in dominant strategies to avoid assumptions about individual beliefs.[2]

Our model of taxation is a simple version of an optimal nonlinear income tax system that is due to Weymark (1987, 1986). Our work is also related to a literature that analyzes public goods under the assumption that a distortionary tax system is used to cover the costs; see, for example, Atkinson and Stern (1974), Boadway and Keen (1993), or Gahvari (2006). This literature, however, is based on a complete information environment in which the distribution of public goods preferences is common knowledge. Hence, there is no problem of preference elicitation.

This latter problem is analyzed in the literature on public good provision under asymmetric information. Recent contributions to this literature include Hellwig (2003), Norman (2004), or Neeman (2004). In these models, individuals differ only in their tastes for a public good. Neither skill heterogeneity nor a redistributive tax system are involved in the analysis. Moreover, the focus lies on the distortions that arise if participation constraints have to be satisfied. In our paper, individuals do not have veto rights and the state's power to raise taxes in order to finance public goods is taken as given.

The idea that a voting rule can be viewed as a mechanism that aggregates information that is dispersed among many individuals goes back to Condorcet (1785); see also Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997) or Piketty (1999). For a model where individuals differ only in their public goods preferences, Ledyard and Palfrey (2002) observe that an optimal provision of public goods can be achieved

[2]This reasoning can be formalized following the approach of Bergemann and Morris (2005). In our framework, robustness of an equilibrium allocation in the sense of Bergemann and Morris (2005) is equivalent to implementability in dominant strategies.

by means of a voting mechanism if the number of individuals is large. Our analysis shows that this result does no longer hold if we consider an economy where individuals differ also in their productive abilities.

The remainder is organized as follows. In Section 2, we specify the model. In Section 3, we consider the problem of public good provision from an optimal mechanism design perspective. Section 4 contains the analysis of democratic mechanisms. The last section concludes.

2 The Model

The Environment

The economy consists of a continuum of individuals, $I := [0, 1]$. The preferences of individual i are given by the utility function

$$\theta^i Q + u(C) - \frac{Y}{w^i} \ ,$$

where Q denotes the quantity of a non-excludable public good, C is the consumption of private goods or after-tax income, and Y is the individual's contribution to the economy's output or pre-tax income. u is a strictly increasing and strictly concave function. An individual's valuation of the public good depends on a taste parameter θ^i that may take two different values, $\theta^i \in \Theta := \{\theta_L, \theta_H\}$ with $0 \leq \theta_L < \theta_H$.[3] The disutility of productive effort depends on the skill parameter w^i, $w^i \in W := \{w_L, w_H\}$ with $0 \leq w_L < w_H$.[4]

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 (i.i.d.)*. The probability that an individual has a high skill level is denoted as

$$\eta \ := \ \text{Prob}\{w^i = w_H\} \ .$$

The random variables $(\tilde{\theta}^i)_{i \in I}$ are also *i.i.d.*. p denotes the probability that any one individual has a high taste parameter,

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

In addition, we assume that a *law of large numbers (LLN)* applies;[5] that is, almost surely, after the realization of randomness at the individual level, the cross-section

[3] The assumption that there are only two possible taste parameters is important for the tractability of the model. However, neither the results in Section 3 nor the characterization of democratic mechanisms in Proposition 3 depend on this assumption.

[4] The assumption that there are only two possible skill levels is made for ease of exposition. Our main result in Proposition 4 can also be proven if W contains a finite number of possible skill levels or if W is a compact interval.

[5] Postulating a LLN for a continuum of *i.i.d.* random variable may creates a measurability problem. There are, however, modeling approaches which circumvent this problem; see Judd (1985) or Al-Najjar (2004).

distribution of characteristics in the economy coincides with the ex ante probability distribution that governs the randomness at the individual level. Accordingly, the probabilities η and p are interpreted as the fractions of individuals with a high earning ability and a high taste parameter, respectively. The *LLN* also implies that the empirical skill distribution and the empirical taste distribution are independent in the sense that the fraction of high-skilled individuals with a high taste parameter and the fraction of low-skilled individuals with a high taste parameter are both equal to p.

We assume that η 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 parameter p is taken to be a random quantity; i.e., there is uncertainty about the distribution of preferences for the public good. The unknown parameter p is henceforth also referred to as the *state of the economy.*

Optimal Income Taxation

We assume that the decisions on taxation and public good provision are made sequentially. First, the decision on public good provision is taken. This generates a tax revenue requirement of $K(Q)$, where K is an increasing, continuously differentiable, and convex cost function, satisfying $K'(0) = 0$ and $\lim_{x\to\infty} K'(x) = \infty$. Given this revenue requirement, taxes are set in order to maximize utilitarian welfare. At this last stage, we assume that the revenue requirement is treated as exogenous. In particular, the tax-setting authorities can not commit to deviate from an ex post optimal tax policy in order to influence the initial decision on public good provision.

As will become clear, this sequential structure has a very convenient implication: It leads to an indirect utility function (see Lemma 1) which makes it very transparent how both an individual's taste parameter and skill level affect his valuation of the public good.[6] Moreover, this indirect utility function is single-peaked which is important for the analysis of voting mechanisms in Section 4.

The tax-setting authority solves the following optimization problem. Choose a tax function $T : Y \mapsto T(Y)$ in order to maximize utilitarian welfare

$$\eta \left(u(Y_H - T(Y_H)) - \frac{Y_H}{w_H} \right) + (1-\eta) \left(u(Y_L - T(Y_L)) - \frac{Y_L}{w_L} \right)$$

subject to the constraints that individual behavior is utility maximizing given the tax function T, for all $t \in \{L, H\}$,

$$Y_t \quad \in \quad \text{argmax}_Y \quad u(Y - T(Y)) - \frac{Y}{w_t} \ ,$$

and the budget constraint,

$$\eta T(Y_H) + (1-\eta) T(Y_L) \quad \geq \quad K(Q) \ .$$

[6]The sequential structure matters only for the analysis of optimal democratic mechanisms in Section 4. Our results on optimal mechanism design in Section 3 do not depend on whether taxes and expenditures are chosen sequentially or simultaneously. In a companion paper, Bierbrauer and Sahm (2008), we assume that taxes and expenditures are determined simultaneously, and show how this affects the analysis of optimal democratic mechanisms. However, this comes at the cost of a more involved analysis of how individual characteristics translate into preferences over public goods.

The solution to this optimization is well known and we only sketch the derivation.[7] Instead of choosing the function T, the tax setting authority can, without loss of generality, be assumed to choose pre- and after-tax income of individuals directly.[8] Hence, the tax setting authority chooses C_L, Y_L, C_H and Y_H in order to maximize

$$\eta\left(u(C_H) - \tfrac{Y_H}{w_H}\right) + (1-\eta)\left(u(C_L) - \tfrac{Y_L}{w_L}\right)$$

subject to the self-selection constraints,

$$u(C_H) - \frac{Y_H}{w_H} \quad \geq \quad u(C_L) - \frac{Y_L}{w_H} \quad \text{and} \quad u(C_L) - \frac{Y_L}{w_L} \quad \geq \quad u(C_H) - \frac{Y_H}{w_L}\ ,$$

and the feasibility constraint

$$\eta(Y_H - C_H) + (1-\eta)(Y_L - C_L) \quad \geq \quad K(Q)\ .$$

At a solution to this problem, the feasibility constraint is binding. Otherwise, the output requirements of all individuals could be lowered without violating the self-selection constraints. Moreover, the self-selection constraint $u(C_H) - Y_H/w_H \geq u(C_L) - Y_L/w_H$ is binding and the self-selection constraint $u(C_L) - Y_L/w_L \geq u(C_H) - Y_H/w_L$ is slack. This is a consequence of the assumption that u is a strictly concave function. A first best utilitarian allocation is such that marginal utilities of consumption are equal, implying that $C_L = C_H$. Moreover, the economy's output is produced only by the productive individuals, $Y_L = 0$ and $Y_H > 0$. This outcome violates the self-selection constraint for the productive individuals. They would rather choose the consumption level C_L without having to produce any output. Hence, redistribution is limited by a binding self-selection constraint.

Using these observations we can solve for Y_L and Y_H as a function of C_L, C_H and Q,

$$\begin{aligned} Y_L &= K(Q) + \eta C_H + (1-\eta)C_L - \eta w_H(u(C_H) - u(C_L))\ , \\ Y_H &= K(Q) + \eta C_H + (1-\eta)C_L + (1-\eta) w_H(u(C_H) - u(C_L))\ . \end{aligned} \tag{1}$$

These expressions can be substituted into the objective function. Utilitarian welfare is then written as a function of C_L and C_H, where C_L and C_H are characterized by the first order conditions,

$$u'(C_L) \quad = \quad \frac{1}{w_L\left(1 - \frac{1}{\lambda}\left(\frac{1}{w_L} - \frac{1}{w_H}\right)\right)} \quad \text{and} \quad u'(C_H) \quad = \quad \frac{1}{w_H}\ , \tag{2}$$

and

$$\lambda \quad := \quad \frac{\eta}{w_H} + \frac{1-\eta}{w_L}\ .$$

The allocation for the productive individuals is undistorted in the sense that their marginal utility of consumption is equal to the marginal disutility of effort. By contrast, the allocation of less productive individuals is distorted since their marginal utility of consumption is strictly larger then their disutility of effort. Hence, the optimal income

[7] A rigorous solution can be found in Hellwig (2007). The special case of preferences that are quasi-linear in leisure is treated by Weymark (1986, 1987).

[8] This observation is known as the *taxation principle*. See, for instance, Guesnerie (1995).

tax implies a downward distortion in the labor supply of low-skilled individuals.

The first order conditions imply that the after-tax incomes of high- and low-skilled individuals do not depend on public good provision. By contrast, an increase in the revenue requirement $K(Q)$ by Δ implies that the pre-tax incomes also increase by Δ. Taking these implications of optimal income taxation into account, individual utilities can be represented in a reduced form.

Lemma 1 Optimal income taxation, given a level Q of the public good, implies that the utility $U(Q,\theta,w)$ of an individual with characteristics $\theta \in \Theta$ and $w \in W$ equals

$$U(Q,\theta,w) \quad = \quad \theta Q - \frac{K(Q)}{w} + \phi(w) \,,$$

where

$$\phi(w_L) \quad := \quad u(C_L) - \frac{1}{w_L}\Big[\tfrac{1}{2}(C_H + C_L) - \tfrac{1}{2}w_H(u(C_H) - u(C_L))\Big] \,,$$

$$\phi(w_H) \quad := \quad u(C_H) - \frac{1}{w_H}\Big[\tfrac{1}{2}(C_H + C_L) + \tfrac{1}{2}w_H(u(C_H) - u(C_L))\Big] \,,$$

and C_L and C_H are given by the first order conditions in (2).

The optimal decision on public good provision

Prior to income taxation, the decision on public good provision has to be taken. For a given p, an optimal choice of Q maximizes

$$\eta\left[p\left(\theta_H Q - \frac{K(Q)}{w_L}\right) + (1-p)\left(\theta_L Q - \frac{K(Q)}{w_L}\right)\right]$$

$$+(1-\eta)\left[p\left(\theta_H Q - \frac{K(Q)}{w_H}\right) + (1-p)\left(\theta_L Q - \frac{K(Q)}{w_H}\right)\right]$$

or, equivalently,

$$\Big(p\theta_H + (1-p)\theta_L\Big)Q - \lambda K(Q) \,.$$

The optimal level of public good provision, $Q^*(p)$, is characterized by the first order condition

$$\lambda K'(Q^*(p)) \quad = \quad p\theta_H + (1-p)\theta_L \,.$$

This optimality condition is a modified version of the Samuelson Rule that takes into account that a distortionary income tax system is used to cover the cost of public good provision. The marginal social cost of public good provision, $\lambda K'(Q^*(p))$, has to be equal to the marginal social benefit, $p\theta_H + (1-p)\theta_L$.

The problem of information aggregation

The function $Q^* : p \mapsto Q^*(p)$ specifies the optimal quantity of the public good as a function of p, i.e., as a function of the cross-section distribution of public goods preferences. If a benevolent planner wants to provide public goods according to this provision rule, she needs to acquire information on the actual value of p. However, individuals have private information about their public goods preferences. This raises the question whether information on p can indeed be obtained if individuals know that the planner seeks to implement provision rule Q^*.

To motivate this problem, suppose that public goods are provided according to Q^*. An individual's utility can then be written as a function of p,

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

With a slight abuse of notation, we drop the term $\phi(w)$ because it does not depend on Q and is hence irrelevant for the problem of preference elicitation. It is easily verified that

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

where

$$v(p) \quad := \quad \frac{p\theta_H + (1-p)\theta_L}{\lambda} \quad = \quad K'(Q^*(p))$$

is a measure of the marginal social benefit from public good provision that takes the marginal cost of public funds, λ, into account. Analogously, θw is a measure of an individual's valuation of the public good that takes into account that public good provision affects the output requirements Y_L and Y_H.

An individual is better off if p is larger – or, equivalently, if more of the public good is provided – if and only if her valuation, θw, exceeds the utilitarian valuation, $v(p)$. Likewise, an individual with a below average valuation prefers to have a lower quantity of the public good. Thus, individuals care about the planner's perception of p. Generally, individuals who pay taxes will not communicate their preferences truthfully if they can influence the planner's decision. Consider an individual with a low taste parameter and a high skill level. Moreover, for the sake of the argument, assume that this individual believes p to be very low.[9] If a vast majority of individuals has a low taste parameter, then this individual can be sure that his own valuation lies above the average, $\theta_L w_H > v(p)$. Put differently, the individual expects that the quantity of the public good is too low. Hence, this individual is inclined to announce a high taste parameter.

We discuss two different approaches to the problem of information aggregation. In Section 3 we discuss the problem from a mechanism design perspective and show that

[9]When this individual decides ex interim what taste parameter to announce, her prior beliefs put a lot of probability mass on values of p that are close to zero.

there is a mechanism that allows to achieve the modified Samuelson Rule. In Section 4 we study a "democratic" environment where tax payers vote over the level of public good provision. In this setting, the modified Samuelson Rule can be no longer reached and we characterize the optimal "democratic" provision rule.

3 Optimal Mechanism Design

In this section, we characterize provision rules that are achievable if individuals are privately informed about their public goods preferences and their skills; that is, we focus on incentive compatibility constraints. The main result is that there is a mechanism that makes it possible to provide public goods according to the modified Samuelson Rule, Q^*. However, this requires that the individuals who provide the information on the distribution of preferences are separated from the set of individuals who cover the cost of public good provision via their income tax payments.

Equilibrium Selection via Sampling

Mechanism design in a continuum economy leads to a severe problem of equilibrium multiplicity. Suppose that public good provision is based on a direct mechanism, i.e., individuals announce their taste parameters and their skills to a mechanism designer. Further, assume that the mechanism designer's decision on public good provision is a function of the measure of individuals who announce a high taste parameter. With such a mechanism, any individual is willing to reveal the own taste parameter because, in a continuum economy, a single individual's announcement does not affect the measure of individuals who announce a high taste parameter. This leads to the trivial conclusion that Q^* is an implementable provision rule because no individual has an impact on public good provision and hence no individual minds revealing her characteristics to the mechanism designer.

However, by the same reasoning, any behavior could be rationalized. For instance, individuals with a low taste parameter and high skills might announce a high taste parameter, because they expect that the mechanism designer will choose a provision level that is too low. More generally, whatever an individual announces, the announcement is a best response because it is inconsequential in a continuum economy.

To deal with this problem of equilibrium multiplicity, we study mechanisms that rely on sampling; that is, a random sample of N individuals is asked to communicate their characteristics to the mechanism designer. Based on the messages of sample members, the mechanism designer decides on public good provision. Finally, individuals who have not been in the sample choose their labor supply subject to an optimal income tax. We study the properties of these sample mechanisms as $N \to \infty$ and an individual's impact on public good provision vanishes.

We invoke the Revelation Principle and limit attention to incentive-compatible, direct mechanisms. Denote the set of sampled individuals by $S_N = \{1, \ldots, N\}$. A direct mechanism consists of a collection of functions $t^i_N : (\Theta \times W)^N \to \mathbb{R}_+$, $C^i_N : (\Theta \times W)^N \to \mathbb{R}_+$, $i \in S_N$, and $Q_N : (\Theta \times W)^N \to \mathbb{R}_+$. $t^i_N((\hat{\theta}^i, \hat{w}^i)_{i\in S_N})$ is an output requirement for sample member i if the profile of announcements to the mechanism is given by $(\hat{\theta}^i, \hat{w}^i)_{i\in S_N}$. Likewise, $C^i_N((\hat{\theta}^i, \hat{w}^i)_{i\in S_N})$ is individual i's consumption as a function of the sample member's announcements. These output requirements and consumption levels are specific to the set of sampled individuals. Finally, $Q_N((\hat{\theta}^i, \hat{w}^i)_{i\in S_N})$ is the decision on public good provision.

A direct mechanism is incentive-compatible, if truth-telling is a dominant strategy:[10] for all $i \in S_N$, for all $(\hat{\theta}^j, \hat{w}^j)_{j\neq i} \in (\Theta \times W)^{N-1}$ and for all $(\theta^i, w^i) \in \Theta \times W$,

$$\theta^i Q_N((\hat{\theta}^j, \hat{w}^j)_{j\neq i}, (\theta^i, w^i)) + u(C^i_N((\hat{\theta}^j, \hat{w}^j)_{j\neq i}, (\theta^i, w^i)))$$

$$-\frac{1}{w^i} t^i_N((\hat{\theta}^j, \hat{w}^j)_{j\neq i}, (\theta^i, w^i))$$

$$\geq \theta^i Q_N((\hat{\theta}^j, \hat{w}^j)_{j\neq i}, (\hat{\theta}^i, \hat{w}^i)) + u(C^i_N((\hat{\theta}^j, \hat{w}^j)_{j\neq i}, (\hat{\theta}^i, \hat{w}^i)))$$

$$-\frac{1}{w^i} t^i_N((\hat{\theta}^j, \hat{w}^j)_{j\neq i}, (\hat{\theta}^i, \hat{w}^i)) \ ,$$

for all $(\hat{\theta}^i, \hat{w}^i) \in \Theta \times W$.

The mechanism designer has the possibility to distinguish between individuals who are in the sample and individuals who are not. The latter pay for the public good via the income tax system. The tax revenues generated from the countable set of individuals in the sample have no weight at an aggregate level. The output requirements $\{t^i_N\}_{i\in S_N}$ and the consumption levels $\{C^i_N\}_{i\in S_N}$ are used only to ensure incentive compatibility for individuals in the sample.

Efficient Provision Rules for Sample Mechanisms

The state of the economy p is a random quantity. For simplicity, we impose the following assumption on the prior beliefs of the mechanism designer.[11]

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

Q_N is said to be an efficient provision rule if $Q_N((\theta^i, w^i)_{i\in S_N})$ solves

$$\max_Q \quad E[(p\theta_H + (1-p)\theta_L)Q - \lambda K(Q) \mid (\theta^i, w^i)_{i\in S_N}] \ ,$$

for every $(\theta^i, w^i)_{i\in S_N}$.

[10] We focus on dominant strategies because this implies that the equilibrium of the mechanism is robust with respect to assumptions about the beliefs of individuals concerning the random variable p and the characteristics of individuals in the sample.

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

Lemma 2 Suppose Assumption 1 holds. Let $m := \#\{i \in S_N \mid \theta^i = \theta_H\}$. Q_N is an efficient provision rule for a sample mechanism if and only if, for every $(\theta^i, w^i)_{i \in S_N}$, $Q_N((\theta^i, w^i)_{i \in S_N})$ equals $Q^* \left(\frac{m+1}{N+2}\right)$. Put differently, any efficient $Q_N((\theta^i, w^i)_{i \in S_N})$ maximizes

$$EW(m, Q) \quad := \quad \left(\frac{m+1}{N+2}\theta_H + \frac{N-m+1}{N+2}\theta_L\right) Q - \lambda K(Q) \ .$$

A mechanism designer has posterior beliefs on p that depend on the number m of individuals in the sample with a high taste parameter. Given these posterior beliefs, the mechanism designer's valuation of the public good is given by

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

In particular, this valuation is an increasing function of the number m of individuals with a high taste parameter.

Proposition 1 For N sufficiently large, an efficient provision rule is incentive-compatible only if output requirements and consumption levels of individuals in the sample are different from those that these individuals would choose under the income tax system.

A proof of Proposition 1 can be found in the Appendix. According to Proposition 1, if the incentives of individuals are shaped only by the income tax, then an efficient use of the information that they provide is not possible. Efficiency can be achieved only via an adjustment of output requirements and consumption levels.
Under an optimal income tax public goods preferences depend on taste parameters and skills. This interaction of taste and skill parameters implies that an income tax does not provide appropriate incentives for a revelation of public goods preferences. To illustrate this, suppose that the possible valuations of the public good are ordered as follows,

$$\theta_L w_L < \theta_H w_L < \theta_L w_H < \theta_H w_H \ .$$

Efficiency requires that if an individual announces a high taste parameter, then more of the public good is provided. Incentive compatibility requires that this outcome makes an individual with a high taste parameter better off. However, if $\theta_L w_H > \theta_H w_L$ and an announcement of θ_H is beneficial for a low skilled individual with a high taste parameter, then it is also beneficial for a high-skilled individual with a low taste parameter; i.e., if incentive compatibility is ensured for all individuals with a high taste parameter, this implies that incentive compatibility is violated for some individuals with a low taste parameter.
The following mechanism is incentive-compatible and achieves an efficient use of information. Individuals in the sample have to produce a fixed amount of output $\bar{Y}$ that

is independent of their announcements.[12] The mechanism designer chooses consumption levels $\{C^L(m)\}_{m=0}^{N-1}$ and $\{C^H(m)\}_{m=1}^{N}$ such that, if there are m individuals with a high taste parameter in the sample, then every individual with a low taste parameter consumes $C^L(m)$ and every individual with a high taste parameter consumes $C^H(m)$, irrespective of the announced skill levels.
A truthful announcement of taste parameters is a dominant strategy if, for every $m \in \{0, \dots, N-1\}$,

$$\theta_L Q_N^*(m) + u(C^L(m)) \quad \geq \quad \theta_L Q_N^*(m+1) + u(C^L(m+1))$$

and

$$\theta_H Q_N^*(m+1) + u(C^H(m+1)) \quad \geq \quad \theta_H Q_N^*(m) + u(C^L(m)) \; .$$

For instance, truth-telling is a (strictly) best response for each individual in the sample if, for every m, $C^L(m)$ and $C^H(m+1)$ are chosen such that

$$u(C^L(m)) - u(C^H(m+1)) \quad = \quad \frac{\theta_H - \theta_L}{2}\Big(Q_N^*(m+1) - Q_N^*(m)\Big) \; .$$

These considerations show that, for any given N, there exists an incentive compatible mechanism that achieves an efficient use of the information that the sample members provide. The following Proposition asserts that for $N \to \infty$ the expected welfare level that is generated by such a mechanism converges to the level of expected welfare that is induced by the Samuelson rule under conditions of complete information on the distribution of preferences. A proof can be found in the Appendix.

Proposition 2 For every N, there exists an incentive compatible mechanism that achieves an efficient use of information. Let EW_N^* be the induced expected welfare level. Then

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

where

$$EW^* \quad = \quad \int_0^1 [(p\theta_H + (1-p)\theta_L)Q^*(p) - \lambda K(Q^*(p))]dp \; ,$$

is the expected welfare level induced by Q^*.

To sum up, under an income tax system, an individual's contribution to a public good is a function of the individual's income. Moreover, an individual's behavior on the labor market depends only on the individual's skills but not on the public goods preferences. Proposition 1 shows that this implies that individuals whose incentives are shaped by the tax system will not reveal their taste parameters. By contrast, a mechanism where

[12] This output requirement does not affect individual incentives. $\bar{Y}$ can, for instance, be set such that, in expectation, the output of sample members is equal to their consumption. Alternatively, $\bar{Y}$ could be set such that participation in the sample is voluntary. Since the sample has only countably many individuals, $\bar{Y}$ does not affect the economy's resource constraint.

a subset of individuals is taken away from the labor market and faces an alternative incentive scheme makes it possible to acquire information on public goods preferences in an efficient way. As Proposition 2 shows, this implies that public goods can be provided as if there was complete information on the distribution of public goods preferences.

4 Democratic Mechanisms

In this section we study democratic mechanisms for income taxation and public good provision. A democratic mechanism is defined by two principles.

The first principle is that individuals vote over public goods. Formally, individuals have an action set with two elements *yes* and *no*. The decision on public good provision is a function of the number of individuals who vote *yes*. If the public good comes as an indivisible unit, $Q \in \{0, 1\}$, then the public good is provided if the number of *yes* votes exceeds some threshold. Our approach is more general in that we consider a public good that can be provided in any quantity. The decision on public good provision is then an increasing function of the number of individuals who vote *yes*; i.e. if more voters support public good provision then a larger quantity of the public good is provided. Conceptually, the difference between a direct mechanism and a voting mechanism is that the former asks individuals for their utility functions. A voting mechanism is ignorant about utility functions, each person – as opposed to each utility function – is given equal consideration for the decision on public good provision.

Our formalization of a voting mechanism has been inspired by the Downes (1957) model of two-party competition. In this model, there are two competing parties. Each party proposes a public goods provision level and the winning party's proposal is finally implemented. Our approach is similar in that (i) we also assume that, a priori, a public good may be provided in any quantity, $Q \in \mathbb{R}_+$, (ii) the decision which quantity is implemented is based on a game in which voters have two actions (party A versus party B in the Downsian model; more versus less public goods provision in our model) and which (iii) has an equilibrium in strategies that are weakly dominant.[13] The key difference between the Downsian model and our approach, of course, is that we do not use political competition to close our model. Instead, given the voting procedure of the Downsian model, we ask what is the best allocation of public goods that one can get. This approach makes it possible to separate the inefficiencies that are due to the use of a voting mechanism from those that are due to political competition.

The second principle is "no taxation without representation". We study the implications of the postulate that a decision on public good provision should reflect the preferences of those individuals who pay taxes. We thus assume that the individuals who decide on public good provision face the same tax system as any other individual in the economy. Formally, the utility function of a voter with characteristics (θ, w) is

[13] In the Downsian model this follows because sincere voting is the only undominated strategy of voters. In our approach, this follows because we seek implementation as a dominant strategy equilibrium.

given by

$$U(Q, \theta, w) \quad = \quad \theta Q - \frac{K(Q)}{w} \ ,$$

that is, by the utility function that is implied by an optimal income tax system.

4.1 Implementation by a voting mechanism

We consider a random sample of N individuals who decide on public good provision via a voting procedure. We will let the sample size go to infinity to single out the "reasonable" equilibrium in the continuum economy. As in the previous section on optimal mechanism design, a provision rule is a function $Q_N : (\Theta \times W)^N \to \mathbb{R}_+$ that specifies how much of the public good is provided as a function of the characteristics of individuals in the sample. In the following we will first formalize the notion that such a provision rule is implementable by a voting mechanism. We will then give a complete characterization of the provision rule with this property.

A voting mechanism is a game where individuals have an action set consisting of the elements *yes* and *no*. The outcome of the voting game is a non-decreasing function $Q_N^V : \{0, \ldots, N\} \to \mathbb{R}_+$ that specifies a decision on public good provision as a function of the number m_y of individuals who vote *yes*.

A strategy $\sigma : \Theta \times W \to \{yes, no\}$ for the game induced by the voting mechanism specifies an individual's vote $\sigma(\theta^i, w^i)$ as a function of the individual's taste parameter and skill level. σ is a dominant strategy if, $\sigma(\theta, w) = no$ implies that

$$U(Q_N^V(m_y), \theta, w) \quad \geq \quad U(Q_N^V(m_y + 1), \theta, w) \ ,$$

for every m_y, and $\sigma(\theta, w) = yes$ implies that

$$U(Q_N^V(m_y), \theta, w) \quad \geq \quad U(Q_N^V(m_y - 1), \theta, w) \ ,$$

for every m_y.

A provision rule $Q_N : (\Theta \times W)^N \to \mathbb{R}_+$ is said to be *implementable by a voting mechanism* if there is a voting mechanism $Q_N^V : \{0, \ldots, N\} \to \mathbb{R}_+$ with an equilibrium σ such that, for every $(\theta^i, w^i)_{i \in S_N}$,

$$Q_N((\theta^i, w^i)_{i \in S_N}) \quad = \quad Q_N^V(m_y^\sigma((\theta^i, w^i)_{i \in S_N})) \ ,$$

where $m_y^\sigma((\theta^i, w^i)_{i \in S_N})$ is the number of yes-votes induced by σ if the characteristics of individuals in the sample are given by $(\theta^i, w^i)_{i \in S_N}$.

Proposition 3 Provision rule Q_N is implementable by a voting mechanism if and only if there is a subset X of the set of types $\Theta \times W$ such that the following properties hold:

i) Whenever (θ, w) belongs to X and (θ', w') does not belong to X then $\theta w \geq \theta' w'$.

ii) Whenever two samples S_N and S'_N are such that $\#\{i \in S_N \mid (\theta^i, w^i) \in X\} = \#\{i \in S'_N \mid (\theta^i, w^i) \in X\}$, then $Q_N((\theta^i, w^i)_{i \in S_N}) = Q_N((\theta^i, w^i)_{i \in S'_N})$. With some abuse of notation we write $Q_N(m_X)$ instead of $Q_N((\theta^i, w^i)_{i \in S_N})$, where $m_X := \#\{i \in S_N \mid (\theta^i, w^i) \in X\}$.

iii) For all $m_X \in \{0, \ldots, N\}$, $(\theta, w) \in X$ implies that

$$U(Q_N(m_X), \theta, w) \quad \geq \quad U(Q_N(m_X - 1), \theta, w) \ , \tag{3}$$

and $(\theta, w) \in -X$ implies that

$$U(Q_N(m_X), \theta, w) \quad \geq \quad U(Q_N(m_X + 1), \theta, w) \ . \tag{4}$$

Proposition 3 gives a complete characterization of the provision rules that can be decentralized via a voting mechanism. It is based on a binary partition of the set of types. The set X contains the individuals who benefit if more of a public good is provided and hence vote *yes*. The complement $-X$ contains the individuals who are harmed if more of the public good is provided and vote *no*.

The proof is in the Appendix. It is based on the observation that any dominant strategy equilibrium of a voting game partitions the set of types $\Theta \times W$ into those who vote *yes*, X, and those who vote *no*, $-X$. Consequently, a decision that reflects the number of individuals who vote *yes* is equivalent to a decision that reflects the number of individuals with types in X.

4.2 The main result

We use the following approach to characterize the optimal voting mechanism. For each subset X of the set of types $\Theta \times W$ that satisfies property i) in Proposition 3, we solve for the provision rule $Q_N : m_X \mapsto Q_N(m_X)$ that maximizes

$$EW_{N,X} \quad := \quad \textstyle\sum_{m_X=0}^{N} \rho(m_X) E[(p\theta_H + (1-p)\theta_L)Q - \lambda K(Q) \mid m_X] \ ,$$

where $\rho(m_X)$ is the probability that the number of sample members with characteristics in X equals m_X, subject to the equilibrium conditions in (3) and (4). We denote the optimal provision rule by $Q^{**}_{N,X}$ and the corresponding level of welfare by $EW^{**}_{N,X}$. We then let $N \to \infty$ and compute $\lim_{N\to\infty} EW^{**}_{N,X}$. Finally we compare the optima that correspond to different values of X to determine the optimal partition, X^*, and thereby the optimal voting mechanism.

Proposition 4

i) The optimal democratic mechanism cannot approximate an efficient decision on public good provision. Formally, $\lim_{N\to\infty} EW^{**}_{N,X^*} < EW^*$.

ii) The discrepancy between the optimal democratic mechanism and an efficient mechanism vanishes as the heterogeneity in abilities vanishes: $\lim_{N\to\infty} EW^{**}_{N,X^*}$ converges to EW^* as w_H converges to w_L.

Theorem 4 follows from Lemmas 3–5 that are discussed in section 4.3. The main statement is that, under a democratic mechanism, heterogeneity in skills is harmful for

public good provision.
To give an intuition, suppose that all individuals have the same skill level. This implies that all individuals with a high taste parameter have an above average valuation of the public good and all individuals with a low taste parameter have a below average valuation of the public good. If public goods are provided according to the Samuelson Rule, then individuals with a high taste parameter vote *yes* to maximize the level of public good provision. Likewise, individuals with a low taste parameter vote *no*. As a consequence, the population share of individuals who vote *yes* is equal to the population share of individuals with a high taste parameter p. Hence, a voting mechanism can achieve public good provision according to the Samuelson Rule.
This reasoning breaks down if there is skill heterogeneity. If an individual with skill level w_L and taste parameter θ_H votes *yes*, then an individual with skill level w_H and taste parameter θ_L will also vote *yes* since the valuation of the public good by the latter, $w_H\theta_L$, is close to or even exceeds $w_L\theta_H$. As a consequence, the provision rule for the public good has to be distorted to ensure that the distribution of votes is informative about the state of the economy. These distortions do not disappear as N goes to infinity.

4.3 Proof of Proposition 4

Without loss of generality, we limit attention to voting mechanisms where the set X of individuals who vote *yes* belongs to

$$\mathfrak{X} := \Big\{\{(\theta_H, w_H)\}, \{(\theta_H, w_L), (\theta_H, w_H)\}, \{(\theta_L, w_H), (\theta_H, w_L), (\theta_H, w_H)\}\Big\} .$$

To see that these are the only cases of interest, consider the partitions of the set of types into a "yes"-set and a "no"-set that are not included in $\mathfrak{X}$. Suppose first that $\theta_L w_H \leq \theta_H w_L$. In this case the only alternatives are voting mechanisms where all individuals vote *no* or all individuals vote *yes*. Clearly, such a voting mechanism does not generate any information about the distribution of public goods preferences.
Now assume that $\theta_L w_H > \theta_H w_L$. In this case, a voting mechanism where individuals with types in $\{(\theta_H, w_L), (\theta_H, w_H)\}$ vote *yes* is not admissible because it violates property i) of Proposition 3. A voting mechanism where all individuals with types in $\{(\theta_L, w_H), (\theta_H, w_H)\}$ vote *yes* would be admissible. However, such a voting mechanism generates information about the number of high-skilled individuals in the sample. Given the assumption that, for every p, skill levels and taste parameters of individuals are independent random variables, this does not generate any information about public goods preferences.
By contrast, the partitions which belong to $\mathfrak{X}$ generate accurate information about p as N goes out of bounds. Intuitively, suppose that $X = \{(\theta_H, w_H)\}$. As N becomes large the number of yes votes converges in probability to $\eta\, p$. Knowing this, the mechanism designer, can infer the distribution of public goods preferences, p, from the number of individuals with a taste parameter and a high skill level, provided that the sample is sufficiently large.

This intuition is formally confirmed by Lemma 3. Before we state this Lemma, we introduce some notation.

Definition 1 For any given $X \in \mathfrak{X}$ denote by Q_X^{**} the maximizer and by EW_X^{**} the maximal value of

$$\int_0^1 [(p\theta_H + (1-p)\theta_L)Q(p) - \lambda K(Q(p))]dp ,$$

where $Q : [0,1] \to \mathbb{R}_+$ has to be chosen such that, for every pair p and p' with $p' > p$, $(\theta, w) \in X$ implies that

$$U(Q(p'), \theta, w) \quad \geq \quad U(Q(p), \theta, w) , \tag{5}$$

and $(\theta, w) \in -X$ implies that

$$U(Q(p), \theta, w) \quad \geq \quad U(Q(p'), \theta, w) . \tag{6}$$

Lemma 3 For any given $X \in \mathfrak{X}$, $\lim_{N\to\infty} EW_{N,X}^{**} = EW_X^{**}$.

A formal proof of Lemma 3 can be found in the Appendix. According to the Lemma, as the sample size goes to infinity, the mechanism designer chooses a provision rule $Q \mapsto Q(p)$ that specifies a decision on public good provision for each state p of the economy. To ensure that this provision rule can be decentralized by a voting mechanism it must be true that all individuals with characteristics in X prefer a large provision level $Q(p')$ over a small provision level $Q(p)$. All other individuals prefer the small provision level over the large provision level.

Lemma 4 Let $w_H > w_L$. Then $EW_X^{**} < EW^*$ for all $X \in \mathfrak{X}$.

Lemmas 4 and 3 imply that statement i) of Proposition 4 is true. We omit a formal proof of Lemma 4 and only sketch the main arguments. Suppose first that w_H is sufficiently close to w_L, so that we have an order of provision levels that is illustrated by the following line,

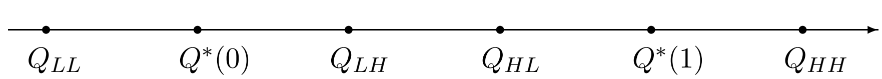

Figure 1. The order of provision levels if $\theta_L w_H < \theta_H w_L$.

where $Q^*(0)$ and $Q^*(1)$ are the minimal and the maximal element, respectively, of the image of the modified Samuelson Rule, Q^*, and Q_{jk} is the maximizer of $U(Q, \theta_j, w_k)$ for $j, k \in \{L, H\}$.

Consider a voting mechanism where all individuals with a high taste parameter vote *yes*. Consequently, for any pair $Q(p')$ and $Q(p)$ with $p' > p$ it must be true that

$$U(Q(p'), \theta_H, w_L) \quad \geq \quad U(Q(p), \theta_H, w_L) \; .$$

This implies that there is at most one p such that $Q(p) \in [Q_{HL}, Q^*(1)]$. Suppose to the contrary that there are $Q(p)$ and $Q(p')$ with $Q_{HL} < Q(p) < Q(p')$. Since $U(Q, \theta_H, w_L)$ is a single-peaked function of Q, $U(Q(p'), \theta_H, w_L) < U(Q(p), \theta_H, w_L)$. Hence, a contradiction. By the same logic, there is at most one p such that $Q(p) \in [Q^*(0), Q_{LH}]$. Otherwise we would contradict $U(Q(p), \theta_L, w_H) \geq U(Q(p'), \theta_L, w_H)$.
Since the image of Q^* contains the intervals $[Q^*(0), Q_{LH}]$ and $[Q_{HL}, Q^*(1)]$, this implies that EW^* cannot be approximated by a voting mechanism where all individuals with a high taste parameter vote *yes*.
The same is true for any alternative voting mechanism. Consider for instance, a voting mechanism where individuals with characteristics in $\{(\theta_L, w_H), (\theta_H, w_L), (\theta_H, w_H)\}$ vote *yes*. Under such a voting mechanism there can be at most one p such that $Q(p) \in [Q_{LH}, Q^*(1)]$. Again, EW^* is out of reach.

Lemma 5 $\max_{X \in \mathfrak{X}} \left\{ EW_X^{**} \right\}$ converges to EW^* as w_H converges to w_L.

Lemmas 5 and 3 imply that statement ii) of Proposition 4 is true. Again, we only sketch the proof. It follows from the observation that Q_{LH} converges to Q_{LL} and Q_{HL} converges to Q_{HH} as w_H converges to w_L. Hence, under a voting mechanism where all individuals with a high taste parameter vote *yes* the distorted intervals $[Q^*(0), Q_{LH}]$ and $[Q_{HL}, Q^*(1)]$ get smaller and smaller. In the limit, the modified Samuelson Rule Q^* is achievable.

4.4 Comparative Statics of Optimal Democratic Mechanisms

We now turn to a characterization of the optimal democratic mechanism. We are particularly interested in the question how the extent of skill heterogeneity affects public good provision under an optimal democratic mechanism.
The analysis proceeds as follows. For each given $X \in \mathfrak{X}$ we characterize the solution to the optimization problem in Definition 1.[14]. For given parameters of the model, this gives a list of candidate solutions that can be used to single out the optimal provision rule and to obtain comparative statics results. While in principle, the comparative statics analysis could be carried out analytically, we illustrate the main insights using a numerical example, in order to avoid lengthy algebraic manipulations. In particular, we assume that $w_H = 1 + x$ and that $w_L = 1 - x$, so that $x \in [0, 1]$ is our measure of skill heterogeneity.[15] Our main findings are the following:

[14]See section B.1 of the Appendix.
[15]A complete description of the example is in section B.2 of the Appendix.

i) Expected welfare is, by and large, a decreasing function of x. However, locally, expected welfare may be increasing in x.

ii) The optimal provision rule for public goods responds less to changes in the distribution of public goods preferences as x increases.

iii) Under an optimal democratic mechanism, the set of types X who are in favor of increased public spending depends on x.

These findings are robust across alternative model specifications and could also be established analytically.[16] In the following we illustrate them by means of our numerical example.

If w_H is sufficiently close to w_L, then the optimal provision rule is such that all individuals with a high taste parameter vote for more public spending, i.e., $X = \{(\theta_H, w_L), (\theta_H, w_H)\}$. Moreover, for x close to zero, the optimal provision rule for the public good has *four bunching regions and an increasing region for medium levels of* p, as in Figure 2 below.

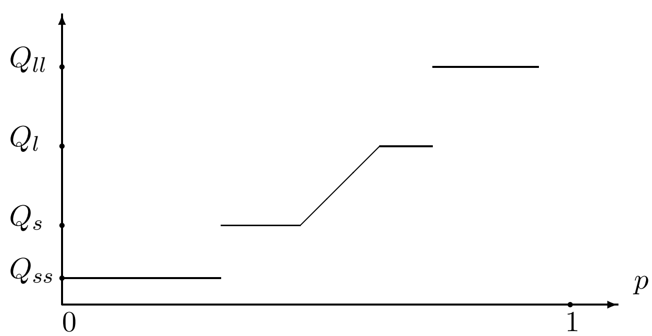

Figure 2. The figure depicts a provision rule with four bunching regions. Over an intermediate range the provision rule coincides with the modified Samuelson Rule, Q^*.

It is necessary to deviate from Q^* for low values of p, because otherwise high-skilled individuals with a low taste parameter vote *yes*, and for high values of p, because otherwise low-skilled individuals with a high taste parameter vote *no*. Formally, for p small, the constraint $U(Q(p-\epsilon), \theta_L, w_H) \geq U(Q(p), \theta_L, w_H)$ is binding. Hence, there is only a choice between Q_{ss} and Q_s, where Q_s is the provision level exceeding Q_{LH} with the property that an individual with type (θ_L, w_H) is indifferent between Q_{ss} and Q_s. For high values of p, the constraint $U(Q(p-\epsilon), \theta_H, w_L) \leq U(Q(p), \theta_H, w_L)$ is binding. This implies that for high levels of p, there is only a choice between the provision levels Q_l and Q_{ll}.

Under an optimal provision rule, Q_s and Q_l get closer to each other as x increases and $\theta_L w_H$ approaches $\theta_H w_L$. A provision rule with *three bunching regions* arises as the limit case in which the range over which the provision rule is continuously increasing shrinks to a single point, see Figure 3.

[16] Details are available from the authors upon request.

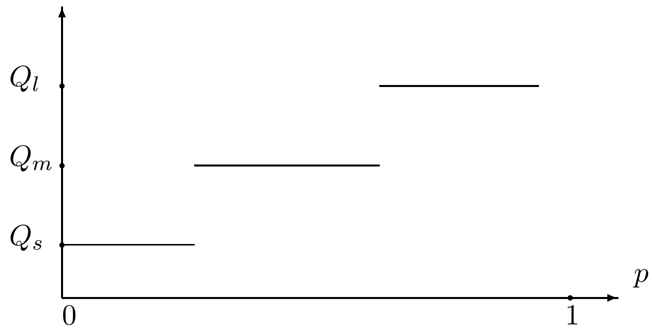

Figure 3. The figure depicts a provision rule with three bunching regions.

As x increases the provision levels Q_s, Q_m and Q_l converge to each other. In the limit there would be a provision rule with only one bunching region, i.e., a provision rule that does not use any information on the state p of the economy. As a consequence, a voting mechanism where all individuals with a high taste parameter vote *yes*, $X = \{(\theta_H, w_l), (\theta_H, w_H)\}$, is not optimal for x sufficiently large. It becomes superior to use a voting mechanism where only individuals with a high taste parameter and a high skill level are in favor of more public spending, $X = \{(\theta_H, w_H)\}$. To illustrate why this is true, consider the scenario in the following figure,

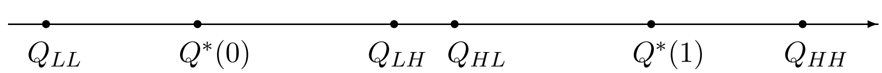

Figure 4. The order of provision levels if $\theta_L w_H$ is close to $\theta_H w_L$.

Compared to a voting mechanism with $X = \{(\theta_H, w_L), (\theta_H, w_H)\}$, a mechanism with $X = \{(\theta_H, w_H)\}$ has the weakness that there can be only one provision level smaller than Q_{HL}, but the strength that the whole interval $[Q_{HL}, Q^*(1)]$ can be used. This advantage is the dominating effect if Q_{HL} is small and the interval $[Q_{HL}, Q^*(1)]$ is large. In this case the optimal provision rule has the shape that is illustrated by Figure 5. This type of provision rule becomes more attractive if x gets larger x and Q_{HL} gets smaller. Consequently, over an intermediate range, $EW^{**}_{X^*}$ is an increasing function of x.

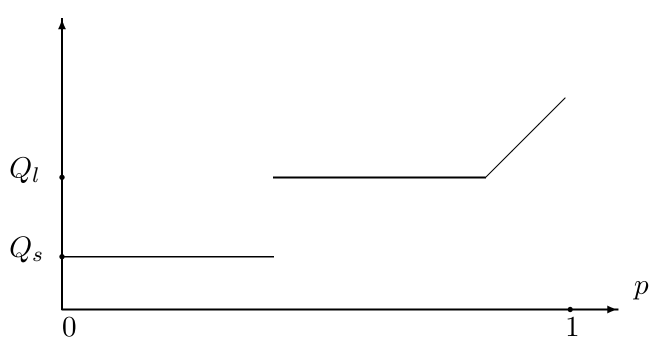

Figure 5. The figure depicts a provision rule with two bunching regions. For large p the

provision rule is strictly increasing and coincides with the modified Samuelson Rule, Q^*.

We eventually reach the point where $\theta_L w_H = \theta_H w_L$, and hence $Q_{LH} = Q_{HL}$. The optimal voting mechanism is still such that $X = \{(\theta_H, w_H)\}$, but now it faces the constraint that there can be only one provision level smaller than Q_{LH}. Again, the shape of the optimal provision rule is as in Figure 5. However, for x and Q_{LH} sufficiently large, the range over which the optimal provision coincides with Q^* shrinks to a singleton. In the limit case $x \to 1$, it yields the same level of welfare as the optimal uninformed provision rule, i.e., the optimal provision rule that requires to choose the same provision level for all p.
The following graph plots the fraction of the surplus

$$FS := \frac{EW^{**}_{X^*} - EW^u}{EW^* - EW^u}$$

that is realized under an optimal democratic mechanism. EW^u denotes the the maximal level of welfare that can be achieved by choosing the same provision level for every p. Hence, $EW^* - EW^u$ is the utility gain from an efficient decision on public good provision and FS is the fraction of this utility gain that is realized under an optimal democratic mechanism.

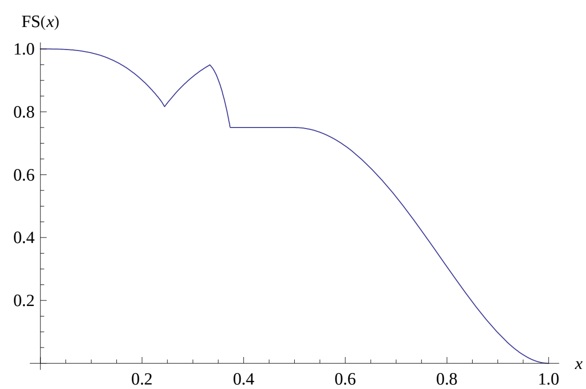

Figure 6. The fraction of the surplus realized by an optimal democratic mechanism as a function of the skill gap x.

The graph shows that FS converges to 1 as x converges to 0 and converges to 0 as x converges to 1. Put differently, if there is no heterogeneity in skills then efficiency is possible under a democratic mechanism. If skill heterogeneity is large, then a democratic mechanism cannot use any information on the distribution of public goods preferences.

5 Concluding Remarks

This paper has studied democratic mechanisms that are defined by the property that a population of tax payers is voting over public goods. It has been shown that these mechanisms perform worse than mechanisms that make use of the possibility to separate the individuals who decide on public good provision from the population of tax payers. By contrast, if all individuals have to pay taxes, then there is an interaction between the tax system and the problem of preference elicitation which implies that inequality is problematic for public good provision. In this sense, our paper provides a critique of democratic mechanisms.

However, if viewed from a different angle, democratic mechanisms are not too bad. Our analysis shows that they are strictly better than "myopic" mechanisms that would be based on the following reasoning: Suppose we take it as given that individuals pay taxes and that the taxes they pay depend on a lot of characteristics such as their abilities. However, the problem at hand is to elicit public goods preferences. So we design a mechanism such that individuals communicate their preferences truthfully and remain ignorant about all the other characteristics. We just make sure that truth-telling is a best response whatever these characteristics are.

In our model we would formalize this as a mechanism where individuals announce a taste parameter and we would require that truth-telling is optimal for every possible skill level. Such a myopic mechanism is equivalent to a voting mechanism where all individuals with high taste parameter are in favor of increased public spending. As our analysis has shown it is in general possible to improve on such a mechanism. For instance, the optimum can be such that only individuals with a high taste parameter and a high skill level vote for more public goods.

We think that, for many applications of mechanism design, a myopic interpretation is not too implausible. The typical textbook treatment is a partial equilibrium analysis where individuals have utility functions that are linear in money. This approach focusses on one particular aspect of heterogeneity and abstracts from all other characteristics that individuals have. If we take this myopic approach as a benchmark, then democratic mechanisms appear quite attractive.

References

Acemoglu, D., Golosov, M., and Tsyvinski, A. (2007). Political economy of mechanisms. MIT Working paper.

Al-Najjar, N. (2004). Aggregation and the law of large numbers in large economies. *Games and Economic Behavior*, 47:1–35.

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

Austen-Smith, D. and Banks, J. (1996). Information aggregation, rationality and the Condorcet jury theorem. *American Political Science Review*, 90:34–45.

Bassetto, M. and Phelan, C. (2008). Tax riots. *Review of Economic Studies*, 75:649–669.

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

Bierbrauer, F. and Sahm, M. (2008). Informative voting, public good provision, and optimal income taxation. Work in Progress.

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

Clarke, E. (1971). Multipart pricing of public goods. *Public Choice*, 11:17–33.

Condorcet, M. (1785). *Essai sur l' application de l'analyse à la probabilité des dècisions rendues à la probabilitè des voix*. Paris: De l'imprimerie royale.

Downes, A. (1957). *An economic theory of democracy*. New York, Harper and Row.

Feddersen, T. and Pesendorfer, W. (1997). Voting behavior and information aggregation in elections with private information. *Econometrica*, 65:1029–1058.

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

Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). Optimal indirect and capital taxation. *Review of Economic Studies*, 70:569–587.

Green, J. and Laffont, J. (1979). *Incentives in Public Decision-Making*. North-Holland Publishing Company.

Groves, T. (1973). Incentives in teams. *Econometrica*, 41:617–663.

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

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

Hellwig, M. (2007). A contribution to the theory of optimal utilitarian income taxation. Preprint 2005/23, Max Planck Institute for Research on Collective Goods, Bonn.

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

Kocherlakota, N. (2005). Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica*, 73:1587–1621.

Ledyard, J. and Palfrey, T. (2002). The approximation of efficient public good mechanisms by simple voting schemes. *Journal of Public Economics*, 83:153–171.

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

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

Norman, P. (2004). Efficient mechanisms for public goods with use exclusion. *Review of Economic Studies*, 71:1163–1188.

Piketty, T. (1999). The information-aggregation approach to political institutions. *European Economic Review*, 43:791–800.

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.

A Appendix

Proof of Lemma 2.

Step 1. We first derive the posterior beliefs of the mechanism designer. Given the assumption that individual taste parameters and skill levels are stochastically independent, it is without loss of generality to assume that the mechanism designer's posterior beliefs are a function of the number ν of agents with a high taste parameter in a sample of size N.

The mechanism designer's prior beliefs are given by the density function ϕ. Under Assumption 1, $\phi(p) = 1$, for all $p \in [0, 1]$. From an ex ante perspective, ν is a random variable, with[17]

$$\begin{aligned} pr(\nu = m) &= \int_0^1 pr(\nu = m \mid p)\phi(p)dp \\ &= \int_0^1 \binom{N}{m} p^m (1-p)^{N-m} dp = \frac{1}{N+1} . \end{aligned} \tag{7}$$

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 thereby induced over p. By Bayes' rule

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

[17]The following relation is used repeatedly:

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

Step 2. A provision level Q induces the following level of expected welfare,

$$\int_0^1 [(p\theta_H + (1-p)\theta_L)Q - \lambda K(Q)]\phi(p \mid \nu = m)dp$$

$$= (N+1)\binom{N}{m}\Big(\int_0^1 [(p\theta_H + (1-p)\theta_L)\, Q - \lambda K(Q)]\, p^m (1-p)^{N-m}\, dp\Big)$$

$$= \left[\frac{m+1}{N+2}\theta_H + \frac{N-m+1}{N+2}\theta_L\right] Q - \lambda K(Q)\ .$$

■

Proof of Proposition 1. Without loss of generality, we assume that preferences of individuals are given in the reduced form $\theta^i Q - \frac{1}{w^i}K(Q)$ that has been derived in Lemma 1.

Step 1. Let $\theta_L w_H \leq \theta_H w_L$. Incentive compatibility of an efficient provision rule holds provided that, for all $m \in \{0, \ldots, N-1\}$, and all $w \in W$,

$$\theta_L Q^*_N(m) - \frac{K(Q^*_N(m))}{w} \quad \geq \quad \theta_L Q^*_N(m+1) - \frac{K(Q^*_N(m+1))}{w}$$

and

$$\theta_H Q^*_N(m+1) - \frac{K(Q^*_N(m+1))}{w} \quad \geq \quad \theta_L Q^*_N(m) - \frac{K(Q^*_N(m))}{w}\ .$$

For N sufficiently large, there exists m such that $Q^*_N(m+1) < Q_{LH}$, where Q_{LH} is the preferred quantity of an individual with $\theta^i = \theta_L$ and $w^i = w_H$, i.e. Q_{LH} is the maximizer of $\theta_L Q - \frac{K(Q)}{w}$. To see this, note that there is $p' \in (0,1)$ such that $Q^*(p') = Q_{LH}$, i.e. there is a state p' such that the modified Samuelson Rule requires to provide the preferred quantity of individuals with $\theta^i = \theta_L$ and $w^i = w_H$. Consequently, $\frac{m+2}{N+2}$ close to zero implies that $Q^*_N(m+1) < Q_{LH}$. Since, the function $\theta_L Q - \frac{K(Q)}{w}$ is single-peaked,

$$\theta_L Q^*_N(m) - \frac{K(Q^*_N(m))}{w_H} \quad < \quad \theta_L Q^*_N(m+1) - \frac{K(Q^*_N(m+1))}{w_H}\ .$$

Hence, a contradiction to incentive compatibility. Analogously, one can show that for $\frac{m}{N+2}$ close to 1 the incentive compatibility constraint for individuals with $\theta^i = \theta_H$ and $w^i = w_L$ is violated.

Step 2. Let $\theta_L w_H > \theta_H w_L$. If the characteristics of i are given by $(\theta^i, w^i) = (\theta_L, w_H)$, then incentive compatibility requires that

$$\theta_L w_H Q^*_N(m) - K(Q^*_N(m)) \quad \geq \quad \theta_L w_H Q^*_N(m+1) - K(Q^*_N(m+1))\ .$$

Analogously, if $(\theta^i, w^i) = (\theta_H, w_L)$, then incentive compatibility requires

$$\theta_H w_L Q^*_N(m) - K(Q^*_N(m)) \quad \leq \quad \theta_H w_L Q^*_N(m+1) - K(Q^*_N(m+1))\ .$$

Adding these inequalities implies that

$$\theta_L w_H \Big(Q^*_N(m+1) - Q^*_N(m)\Big) \quad \geq \quad \theta_H w_L \Big(Q^*_N(m+1) - Q^*_N(m)\Big)\ ,$$

or, equivalently, $\theta_L w_H \geq \theta_H w_L$. Hence, a contradiction. ■

Proof of Proposition 2. Let Q_N be an efficient provision rule for a sample of size N. The induced level of expected welfare is given by

$$\begin{aligned} EW_N^* &= \frac{1}{N+1}\sum_{m=0}^{N} EW\left(m, Q^*\left(\frac{m+1}{N+2}\right)\right) , \\ &= \frac{1}{N+1}\sum_{m=0}^{N}\left(\frac{m+1}{N+2}\theta_H + \frac{N-m+1}{N+2}\theta_L\right) Q^*\left(\frac{m+1}{N+2}\right) - \lambda K\left(Q^*\left(\frac{m+1}{N+2}\right)\right) \end{aligned}$$

This term is known as the *Riemann sum* for the function $(p\theta_H + (1-p)\theta_L)Q^*(p) - \lambda K(Q^*(p))$ and thus converges to EW^* for growing N. ■

Proof of Proposition 3.

"$\Longleftarrow$": Suppose provision rule Q_N satisfies i), ii) and iii). Define a voting mechanism Q_N^V such that for every $z \in \{0, \ldots, N\}$, $Q_N(z) = Q_N^V(z)$. Consider the strategy σ with $\sigma(\theta, w) = yes$, whenever $(\theta, w) \in X$, and $\sigma(\theta, w) = no$, otherwise.

We first show that σ is an equilibrium in dominant strategies for the game induced by Q_N^V. Suppose otherwise, then there is m_y and (θ, w) such that $\sigma(\theta, w) = no$ and $U(Q_N^V(m_y), \theta, w) < U(Q_N^V(m_y - 1), \theta, w)$ or $\sigma(\theta, w) = yes$ and $U(Q_N^V(m_y), \theta, w) < U(Q_N^V(m_y+1), \theta, w)$. But this implies that there is m_X and (θ, w) such that $(\theta, w) \in -X$ and $U(Q_N(m_X), \theta, w) < U(Q_N(m_X - 1), \theta, w)$ or $(\theta, w) \in X$ and $U(Q_N(m_X), \theta, w) < U(Q_N(m_X + 1), \theta, w)$. Hence, a contradiction to the assumption that iii) holds.

It remains to show that $Q_N((\theta^i, w^i)_{i\in S_N}) = Q_N^V(m_y^\sigma((\theta^i, w^i)_{i\in S_N}))$, for every $(\theta^i, w^i)_{i\in S_N}$. This follows from the fact that Q_N can be viewed a function of m_X and that for every m_X, $m_X = m_y^\sigma((\theta^i, w^i)_{i\in S_N})$.

"$\Longrightarrow$": Suppose provision rule Q_N is implementable by a voting mechanism Q_N^V with dominant strategy σ.

We first show that $\sigma(\theta, w) = no$ and $\sigma(\theta', w') = yes$ implies that $\theta w \leq \theta' w'$, for any pair (θ, w) and (θ', w'). Suppose to the contrary that $\theta w > \theta' w'$, then there exists m_y such that

$$\theta' w' Q_N^V(m_y) - K(Q_N^V(m_y)) \quad \geq \quad \theta' w' Q_N^V(m_y - 1) - K(Q_N^V(m_y - 1)) ,$$

and

$$\theta w Q_N^V(m_y - 1) - K(Q_N^V(m_y - 1)) \quad \geq \quad \theta w Q_N^V(m_y) - K(Q_N^V(m_y)) .$$

Adding these constraints implies that

$$(\theta' w' - \theta w)(Q_N^V(m_y) - Q_N^V(m_y - 1)) \quad \geq \quad 0$$

or, since Q_N^V is a non-decreasing function, $\theta' w' \geq \theta w$. Hence a contradiction.

Define the set $X' = \{(\theta, w) \in \Theta \times W \mid \sigma(\theta, w) = yes\}$. We show that Q_N satisfies i) and ii) for X'.

Consider two samples S_N and S_N' such that $\#\{i \in S_N \mid (\theta^i, w^i) \in X'\} = \#\{i \in S_N' \mid$

$(\theta^i, w^i) \in X'\}$. This implies that $m_y^\sigma((\theta^i, w^i)_{i\in S_N}) = m_y^\sigma((\theta^i, w^i)_{i\in S'_N})$. Since Q_N^V implements Q_N, this also implies that $Q_N((\theta^i, w^i)_{i\in S_N}) = Q_N((\theta^i, w^i)_{i\in S'_N})$.
We finally show that iii) holds. Suppose otherwise, then there exists m_X and (θ, w) such that $(\theta, w) \in -X'$ and $U(Q_N(m_X), \theta, w) < U(Q_N(m_X - 1), \theta, w)$ or $(\theta, w) \in X'$ and $U(Q_N(m_X), \theta, w) < U(Q_N(m_X + 1), \theta, w)$. Since Q_N^V implements Q_N, it must be true that $Q_N(z) = Q_N^V(z)$, for all $z \in \{0, \dots, N\}$. Hence, there is m_y and (θ, w) such that $\sigma(\theta, w) = no$ and $U(Q_N^V(m_y), \theta, w) < U(Q_N^V(m_y - 1), \theta, w)$ or $\sigma(\theta, w) = yes$ and $U(Q_N^V(m_y), \theta, w) < U(Q_N^V(m_y + 1), \theta, w)$. But this contradicts the assumption that σ is a dominant strategy. ■

Proof of Lemma 3. The proof consists of two parts. In part A, we show that the statement is true for the case of $X = \{(\theta_H, w_L), (\theta_H, w_H)\}$. In part B, we demonstrate that the same arguments also apply to the case of $X = \{(\theta_H, w_H)\}$ (and analogously to $X = \{(\theta_L, w_H), (\theta_H, w_L), (\theta_H, w_H)\}$).
Part A. Consider $X = \{(\theta_H, w_L), (\theta_H, w_H)\}$.
Step 1. Using the computations in the proof of Lemma 2, we derive

$$EW_{N,X}^{**} = \lambda \frac{1}{N+1} \sum_{m_X=0}^{N} \left\{ v\left(\frac{m_X+1}{N+2}\right) Q_{N,X}^{**}(m_X) - K(Q_{N,X}^{**}(m_X)) \right\}. \tag{8}$$

Step 2. We now show that $EW_{N,X}^{**} \leq EW_X^{**}$ for any $N \in \mathbb{N}$. To establish this claim, we consider, for any given $N \in \mathbb{N}$, the following piecewise constant function $\tilde{Q}_{N,X}^{**} : [0,1] \to \{Q_{N,X}^{**}(m_X)\}_{m_X=0}^N$ with

$$\begin{aligned} \tilde{Q}_{N,X}^{**}(p) &:= Q_{N,X}^{**}(m_X) \quad \text{for} \quad \frac{m_X}{N+1} \leq p < \frac{m_X+1}{N+1},\ m_X \in \{0, \dots, N-1\} \\ \tilde{Q}_{N,X}^{**}(1) &:= Q_{N,X}^{**}(N). \end{aligned}$$

The welfare level induced by $\tilde{Q}_{N,X}^{**}$ is denoted $\widetilde{EW}_{N,X}^{**}$. By definition, $\tilde{Q}_{N,X}^{**}$ is monotonically increasing in p and meets the constraints (5) and (6) since $Q_{N,X}^{**}$ meets (3) and (4). Hence, by the optimality of Q_X^{**} among the provision rules satisfying (5) and (6), $\widetilde{EW}_{N,X}^{**} \leq EW_X^{**}$. It thus suffices to show that $EW_{N,X}^{**} \leq \widetilde{EW}_{N,X}^{**}$.
In order to compute $\widetilde{EW}_{N,X}^{**}$, we first collect a number of observations which are easily verified by the reader.

1. For all $\underline{p}, \bar{p} \in [0,1]$ we have $\int_{\underline{p}}^{\bar{p}} v(p)dp = (\bar{p} - \underline{p}) v\left(\frac{\bar{p}+\underline{p}}{2}\right)$.
2. For all $m \in \{0, 1, \dots, N\}$ we have $\frac{m+\frac{1}{2}}{N+1} = \frac{m+1}{N+2} + \frac{m - \frac{1}{2}N}{(N+1)(N+2)}$.
3. For all $x, y \in [0,1]$ with $x + y \in [0,1]$ we have $v(x+y) = v(x) + \frac{\theta_H - \theta_L}{\lambda} y$.

Using these equalities and step 1, we calculate

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

It remains to show that $\sum_{m_X=0}^{N}(m_X - \frac{1}{2}N)Q^{**}_{N,X}(m_X) \geq 0$. This expression equals

$$\textstyle\sum_{m_X=0}^{\frac{1}{2}N}(\frac{1}{2}N - m_X)(Q^{**}_{N,X}(N - m_X) - Q^{**}_{N,X}(m_X))$$

if N is even and

$$\textstyle\sum_{m_X=0}^{\frac{N-1}{2}}(\frac{1}{2}N - m_X)(Q^{**}_{N,X}(N - m_X) - Q^{**}_{N,X}(m_X))$$

if N is odd. However, as $Q^{**}_{N,X}$ is increasing, those sums are non-negative.

Step 3. We now establish that $EW^{**}_{N,X}$ converges to EW^{**}_X as $N \to \infty$. To this end, for any $N \in \mathbb{N}$, consider the *restriction* $Q^{**}_{X|N}$ of Q^{**}_X to the domain $\{0, 1, \ldots, N\}$. Formally, for each $m_X \in \{0, 1, \ldots, N\}$, set $Q^{**}_{X|N}(m_X) := Q^{**}_X\left(\frac{m_X}{N}\right)$. By definition, $Q^{**}_{X|N}$ is monotonically increasing in m_X for any $N \in \mathbb{N}$ and meets the constraints (3) and (4) as Q^{**}_X satisfies (5) and (6). Denote by $EW^{**}_{X|N}$ the expected welfare level induced by $Q^{**}_{X|N}$. Noting that $Q^{**}_{N,X}$ is optimal among the provision rules satisfying (3) and (4) and using step 2, we then have $EW^{**}_{X|N} \leq EW^{**}_{N,X} \leq EW^{**}_X$ for any $N \in \mathbb{N}$. Thus it suffices to show that $\lim_{N\to\infty} EW^{**}_{X|N} = EW^{**}_X$. To see this, we calculate

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

The first term is the *Riemann sum* for $v(p)Q^{**}_X(p) - K(Q^{**}_X(p))$ and thus converges to EW^{**}_X for growing N. The absolute value of the second term in the sum is bounded from above by the expression $\frac{\theta_H-\theta_L}{N+2}Q^{**}_X(1)$, which vanishes as $N \to \infty$.

Part B. Consider $X = \{(\theta_H, w_H)\}$.[18] Let $\nu_X := \#\{i \in S_N \mid (\theta^i, w^i) = (\theta_H, w_H)\}$ and $\nu := \#\{i \in S_N \mid \theta^i = \theta_H\}$. We show that for large N, the objective function for a voting mechanism with $X = \{(\theta_H, w_H)\}$ has the same mathematical structure as a voting mechanism with $X = \{(\theta_H, w_L), (\theta_H, w_H)\}$. This implies that the same arguments can be used to establish convergence.

Let $Q_{N,X} : m_X \mapsto Q_{N,X}(m_X)$ be an admissible provision rule for a voting mechanism with $X = \{(\theta_H, w_H)\}$. Using similar computations as in the proof of Lemma 2, we derive the induced level of expected welfare

$$\begin{aligned}
&\lambda\sum_{m_X=0}^{N}\mathrm{pr}(\nu_X = m_X)\left[v\left(\frac{1}{\eta}\frac{m_X+1}{N+2}\right)Q_{N,X}(m_X) - K(Q_{N,X}(m_X))\right] \\
&-\frac{1-\eta}{\eta}\frac{\theta_H - \theta_L}{N+2}\sum_{m_X=0}^{N}\binom{N}{m_X}\eta^{m_X}(1-\eta)^{N-m_X}Q_{N,X}(m_X),
\end{aligned} \tag{9}$$

[18] An analogous argument can be made for $X = \{(\theta_L, w_H), (\theta_H, w_L), (\theta_H, w_H)\}$.

where $\mathrm{pr}(\nu_X = m_X) = \frac{1}{N+1}\sum_{k=0}^{N-m_x}\binom{N+1}{m_X}\eta^{N-k}(1-\eta)^k$. The sum in the second line can without loss of generality be assumed to be bounded because it will never be optimal to choose an infinitely large provision level of the public good; hence the term in the second line vanishes as $N \to \infty$.

The random variable $\frac{1}{\eta}\frac{m_X+1}{N+2}$ is a consistent estimator for the unknown value of $p \in [0,1]$.[19] The random variable $\frac{\nu_{X'}}{N}$ for $X' = \{(\theta_H, w_L), (\theta_H, w_H)\}$ is also a consistent estimator of p. Moreover, the probability distributions of both random variables converge to a uniform density as $N \to \infty$.[20] As a consequence, for large N, choosing $Q_{N,X} : m_X \mapsto Q_{N,X}(m_X)$ to maximize

$$\lambda \sum_{m_X=0}^{N} \mathrm{pr}(\tfrac{1}{\eta}\tfrac{\nu_X}{N} = \tfrac{1}{\eta}\tfrac{m_X}{N}) \left[v\left(\frac{1}{\eta}\frac{m_X+1}{N+2}\right) Q_{N,X}(m_X) - K(Q_{N,X}(m_X)) \right]$$

such that $(\theta, w) \in \{(\theta_H, w_H)\}$ implies $U(Q_{N,X}(m_X), \theta, w) \geq U(Q_{N,X}(m_X - 1), \theta, w)$ and that $(\theta, w) \in -\{(\theta_H, w_H)\}$ implies $U(Q_{N,X}(m_X, \theta, w) \geq U(Q_{N,X}(m_X + 1), \theta, w)$, yields the same level of welfare as choosing $Q_{N,X'} : m_{X'} \mapsto Q_{N,X'}(m_{X'})$ to maximize

$$\frac{\lambda}{N+1} \sum_{m_{X'}=0}^{N} \left[v\left(\frac{m_{X'}+1}{N+2}\right) Q_{N,X'}(m_{X'}) - K(Q_{N,X'}(m_{X'})) \right]$$

subject to the constraint that $(\theta, w) \in \{(\theta_H, w_H)\}$ implies $U(Q_{N,X'}(m_{X'}), \theta, w) \geq U(Q_{N,X'}(m_{X'} - 1), \theta, w)$ and that $(\theta, w) \in -\{(\theta_H, w_H)\}$ implies $U(Q_{N,X'}(m_{X'}, \theta, w) \geq U(Q_{N,X'}(m_{X'} + 1), \theta, w)$.

Given the latter optimization problem we can apply the arguments from part A again to establish convergence.

■

[19] To see this note that – given any sample realization where m_X individuals have characteristics in X – the maximum likelihood estimator $\hat{p}_{ML}$ of p satisfies

$$\hat{p}_{ML} = \mathrm{argmax}_{s\in[0,1]}(\eta s)^{m_X}(1-\eta s)^{N-m_X} .$$

Hence $\hat{p}_{ML} = \frac{1}{\eta}\frac{m_x}{N}$. The maximum likelihood estimator is known to be be consistent. This implies that $\frac{1}{\eta}\frac{m_X+1}{N+2}$ is also consistent.

[20] We demonstrate this for $\frac{1}{\eta}\frac{\nu_X}{N}$. Let $0 \leq x < y \leq 1$. We have that

$$\mathrm{pr}(x \leq \tfrac{1}{\eta}\tfrac{\nu_X}{N} \leq y) = \int_0^1 \mathrm{pr}(x \leq \tfrac{1}{\eta}\tfrac{\nu_X}{N} \leq y)dp ,$$

where the consistency of $\frac{1}{\eta}\frac{\nu_X}{N}$ implies that for large N, $\mathrm{pr}(x \leq \frac{1}{\eta}\frac{\nu_X}{N} \leq y \mid p) = 1$ if $x \leq p \leq y$ and $\mathrm{pr}(x \leq \frac{1}{\eta}\frac{\nu_X}{N} \leq y \mid p) = 0$ otherwise. Hence,

$$\mathrm{pr}(x \leq \tfrac{1}{\eta}\tfrac{\nu_X}{N} \leq y) = y - x .$$

Which proves that $\frac{1}{\eta}\frac{\nu_X}{N}$ converges pointwise to a continuous random variable that is uniformly distributed over the interval $[0,1]$.

B Appendix

B.1 A taxonomy of possible solutions

Lemma 6 Consider voting mechanisms where the set of individuals who vote *yes* is given by $X = \{(\theta_H, w_L), (\theta_H, w_H)\}$. An optimal provision rule belongs to one of the following categories:

i) *Provision rules that are constant over four "bunching regions" and that are increasing for medium levels of p.* For such a provision rule there exist numbers Q_{ss}, Q_s, Q_l and Q_{ll} such that

$$Q(p) \quad = \quad \begin{cases} Q_{ss} & \text{for} \quad 0 \leq p \leq \hat{p} \ , \\ Q_s & \text{for} \quad \hat{p} < p < \hat{p}' \ , \\ Q^*(p) & \text{for} \quad \hat{p}' \leq p \leq \tilde{p}' \ , \\ Q_l & \text{for} \quad \tilde{p}' < p < \tilde{p} \ , \\ Q_{ll} & \text{for} \quad \tilde{p} \leq p \leq 1 \ , \end{cases}$$

where $U(Q_{ss}, \theta_L, w_H) = U(Q_s, \theta_L, w_H)$, $U(Q_l, \theta_H, w_L) = U(Q_{ll}, \theta_H, w_L)$ and the critical indices are implicitly defined by the following equations:

$$\bar{v}(\hat{p}) = \theta_L w_H \ , \quad Q^*(\hat{p}') = Q_s \ , \quad Q^*(\tilde{p}') = Q_l \ , \quad \bar{v}(\tilde{p}) = \theta_H w_L \ .$$

ii) *Provision rules that are constant over three "bunching regions".* For such a provision rule there exist numbers Q_s, Q_m and Q_l such that

$$Q(p) \quad = \quad \begin{cases} Q_s & \text{for} \quad 0 \leq p \leq \hat{p} \ , \\ Q_m & \text{for} \quad \hat{p} < p < \tilde{p} \ , \\ Q_l & \text{for} \quad \tilde{p} \leq p \leq 1 \ , \end{cases}$$

where $U(Q_s, \theta_L, w_H) = U(Q_m, \theta_L, w_H)$, $U(Q_m, \theta_H, w_L) = U(Q_l, \theta_H, w_L)$ and the critical indices $\hat{p}$ and $\tilde{p}$ are defined implicitly by the equations

$$\bar{v}(\hat{p}) = \theta_L w_H \quad \text{and} \quad \bar{v}(\tilde{p}) = \theta_H w_L \ .$$

iii) *that are constant over two "bunching regions".* For such a provision rule there exist numbers Q_s and Q_l such that

$$Q(p) \quad = \quad \begin{cases} Q_s & \text{for} \quad 0 \leq p \leq \hat{p} \ , \\ Q_l & \text{for} \quad \hat{p} < p \leq 1 \ , \end{cases}$$

where the critical index $\hat{p}$ is defined implicitly by the equation

$$\bar{v}(\hat{p})Q_s - K(Q_s) \quad = \quad \bar{v}(\hat{p})Q_l - K(Q_l) \ .$$

Proof *Claim 1.* Denote by $\mathcal{V}_Q$ the image of a provision rule $Q : p \mapsto Q(p)$, i.e. $x \in \mathcal{V}_Q$ if and only if there exists $p \in [0,1]$ with $Q(p) = x$. Suppose Q solves the problem of maximizing EW subject to the following informative voting (IV) constraints: for every pair p and p' with $p' > p$, $(\theta, w) \in \{(\theta_H, w_L), (\theta_H, w_H)\}$ implies that $U(Q(p'), \theta, w) \geq U(Q(p), \theta, w)$, and $(\theta, w) \in \{(\theta_L, w_L), (\theta_L, w_H)\}$ implies that $U(Q(p), \theta, w) \geq U(Q(p'), \theta, w)$. Then there exists at most one element $x \in \mathcal{V}_Q$ with $x < Q_{LH}$ and at most one element $z \in \mathcal{V}_Q$ with $z > Q_{HL}$.
Proof of Claim 1. Suppose, to the contrary, that there exist $x, y \in \mathcal{V}_Q$ with $x < y < Q_{LH}$. This implies that there exist p and $p' > p$, with $Q(p) < Q(p') < Q_{LH}$. Since preferences are single-peaked,

$$\theta_L w_H Q(p) - K(Q(p)) < \theta_L w_H Q(p') - K(Q(p')) \ ,$$

a contradiction to the IV constraints. Analogously one shows that the image of an admissible provision rule contains at most one element z with $z > Q_{HL}$.
Claim 2. A provision rule Q for which there exists $y \in \mathcal{V}_Q$ with $y \in [Q_{LH}, Q_{HL}]$ is a candidate for a solution only if there exist $x, z \in \mathcal{V}_Q$ with $x < Q_{LH}$ and $Q_{HL} < z$.
Proof of Claim 2. We first argue that a provision rule Q for which there exist neither $x \in \mathcal{V}_Q$ with $x < Q_{LH}$ nor $z \in \mathcal{V}_Q$ with $z > Q_{HL}$ cannot be optimal. To be optimal, such a hypothetical provision rule would have to be the degenerate case of a provision rule with four pooling levels, which results as the limit outcome as Q_{ss} converges to Q_{LH} and Q_l converges to Q_{HL}. Under a provision rule characterized by four pooling levels, expected welfare EW satisfies the following equation:

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

where Q_s and $\hat{p}'$ are implicit functions of Q_{ss}. Similarly, Q_l and $\tilde{p}'$ are implicit functions of Q_{ll}. Taking these functional relationships into account, one may compute the partial derivatives and verify that

$$\lim_{Q_{ss} \to Q_{LH}} \frac{\partial EW(Q_{ss}, Q_{ll})}{\partial Q_{ss}} < 0 \quad \text{and} \quad \lim_{Q_{ll} \to Q_{HL}} \frac{\partial EW(Q_{ss}, Q_{ll})}{\partial Q_{ll}} > 0.$$

Thus, $Q_{ss} = Q_{LH}$ and $Q_{ll} = Q_{HL}$ cannot be optimal.
We now argue in a similar manner that it cannot be optimal to choose a provision rule with $y, z \in \mathcal{V}_Q$ satisfying $Q_{LH} < y < Q_{HL} < z$, but without $x \in \mathcal{V}_Q$ satisfying $x < Q_{LH}$:
Define $\tilde{z} < Q_{HL}$ by the equation $\theta_H w_L \ z - K(z) = \theta_H w_L \tilde{z} - K(\tilde{z})$. Note that for such a provision rule to be a optimal under it has to be true that $y \leq \tilde{z}$ and that

$\mathcal{V}_Q = [Q_{LH}, \tilde{z}] \cup \{z\}$ by Claim 1. Again, this is a degenerate case of a provision rule with four pooling levels, namely the one that results as Q_{ss} converges to Q_{LH} and $Q_{ll} = z$. As above, this hypothetical solution can be ruled out as

$$\lim_{Q_{ss} \to Q_{LH}} \frac{\partial EW(Q_{ss}, Q_{ll})}{\partial Q_{ss}} \quad < \quad 0 \, .$$

The analogous argument allows us to rule out a provision rule with $x, y \in \mathcal{V}_Q$ and $x < Q_{LH} < y < Q_{HL}$ but without $z \in \mathcal{V}_Q$ satisfying $z > Q_{HL}$.
These arguments imply that an optimal provision rule has to be one with two, three or four bunching regions. ■

Lemma 7 Consider voting mechanisms where the set of individuals who vote *yes* is given by $X = \{(\theta_H, w_H)\}$. An optimal provision rule belongs to one of the following categories:

i) *Provision rules that are constant over two bunching regions and that are increasing for high levels of p.* For such a provision rule there exist numbers Q_s and Q_l such that

$$Q(p) \quad := \quad \begin{cases} Q_s & \text{for} \quad 0 \leq p \leq \hat{p} \, , \\ Q_l & \text{for} \quad \hat{p} < p < \hat{p}' \, , \\ Q^*(p) & \text{for} \quad \hat{p}' \leq p \leq 1 \, . \end{cases}$$

If $\theta_L w_H \leq \theta_H w_L$, then $U(Q_s, \theta_H, w_L) = U(Q_l, \theta_H, w_L)$ and the critical indices are defined by $\bar{v}(\hat{p}) = \theta_H w_L$ and $Q^*(\hat{p}') = Q_l$. If $\theta_L w_H > \theta_H w_L$, then $U(Q_s, \theta_L, w_H) = U(Q_l, \theta_L, w_H)$ and the critical indices are defined by $\bar{v}(\hat{p}) = \theta_L w_H$ and $Q^*(\hat{p'}) = Q_l$.

ii) *Provision rules with two bunching regions.*

Proof The proof uses similar arguments as the proof of Lemma 6. Hence, the arguments are only sketched. Suppose Q solves the problem of maximizing EW subject to the following informative voting (IV) constraints: for every pair p and p' with $p' > p$, $(\theta, w) \in \{(\theta_H, w_H)\}$ implies that $U(Q(p'), \theta, w) \geq U(Q(p), \theta, w)$, and $(\theta, w) \in \{(\theta_H, w_L), (\theta_L, w_L), (\theta_L, w_H)\}$ implies that $U(Q(p), \theta, w) \geq U(Q(p'), \theta, w)$.
Claim. If $\theta_L w_H < \theta_H w_L$, then there is exactly one element $x \in \mathcal{V}_Q$ with $x < Q_{LH}$. If $\theta_H w_L \leq \theta_L w_H < \theta_H w_L$, then there is exactly is at most one element $x \in \mathcal{V}_Q$ with $x < Q_{HL}$.
Proof. Because preferences are single-peaked there can be at most one $x \in \mathcal{V}_Q$ with $x < Q_{LH}$ ($x < Q_{HL}$) if $\theta_L w_H < \theta_H w_L$ ($\theta_H w_L \leq \theta_L w_H$). It follows from the same argument as in Claim 3 in the proof of Lemma 6 that it cannot be optimal to have no $x \in \mathcal{V}_Q$ with $x < Q_{LH}$ ($x < Q_{HL}$) if $\theta_L w_H < \theta_H w_L$ ($\theta_H w_L \leq \theta_L w_H$).
This implies that an optimal provision rule under IV constraints, which is not constant, is constant over two bunching regions. Depending on the parameters it may be

optimal to choose the larger bunching point Q_l such that $Q_l < Q^*(1)$. In this case there is a range of large values of p, where the optimal provision rule coincides with the Samuelson Rule Q^*. These are the levels of p for which $Q^*(p) \geq Q_l$. ■

The analysis of a voting mechanism where only individuals with characteristics given by (θ_L, w_L) vote *no* is the mirror image of the case where only individuals with (θ_H, w_H) vote *yes*. The proof uses the same arguments as the proof of Lemma 7 and is omitted.

Lemma 8 Consider voting mechanisms where the set of individuals who vote *yes* is given by $X = \{(\theta_H, w_L), (\theta_L, w_H), (\theta_H, w_H)\}$. An optimal provision rule belongs to one of the following categories:

i) *Provision rules that are constant over two bunching regions and that are increasing interval for small levels of p.* For such a provision rule there exist numbers Q_s and Q_l such that

$$Q(p) \quad := \quad \begin{cases} Q^*(p) & \text{for} \quad 0 \leq p \leq \hat{p} \ , \\ Q_s & \text{for} \quad \hat{p} < p < \hat{p}' \ , \\ Q_l & \text{for} \quad \hat{p}' \leq p \leq 1 \ . \end{cases}$$

If $\theta_L w_H \leq \theta_H w_L$, then $U(Q_s, \theta_L, w_H) = U(Q_l, \theta_L, w_H)$ and the critical indices are defined by $\bar{v}(\hat{p}') = \theta_L w_H$ and $Q^*(\hat{p}) = Q_s$. If $\theta_L w_H > \theta_H w_L$, then $U(Q_s, \theta_H, w_L) = U(Q_l, \theta_H, w_L)$ and the critical indices are defined by $\bar{v}(\hat{p}') = \theta_H w_L$ and $Q^*(\hat{p}) = Q_l$.

ii) *Provision rules with two bunching intervals.*

B.2 Results of the comparative statics exercise

Our numerical comparative statics exercise is based on the following assumptions.

Assumption 2 Let $\theta_L = 1$, and $\theta_H = 2$. Suppose that $K(Q) = \frac{1}{2}Q^2$. Let $w_L = 1 - x$ and $w_H = 1 + x$. For any x, let $\lambda = 1$.

The assumptions $\theta_L = 1$ and $\theta_H = 2$ are normalizations. Assuming a quadratic cost function is a simplification that implies that the optimal provision level from the perspective of an individual with characteristics (θ, w) equals θw and that this individual prefers a provision level Q' over Q'' if and only if $\mid \theta w - Q' \mid \leq \mid \theta w - Q'' \mid$.
x is our measure of skill heterogeneity. We let x vary between 0 and 1 and assume that the shadow cost of public funds λ remains constant. Hence, as we vary x, we are also adjusting the share of high-skilled individuals in the population, η, so that λ remains unaffected. Holding λ fixed at 1 implies that the modified Samuelson Rule Q^* is given by $Q^*(p) = 1 + p$, so that our first best benchmark remains constant. Put differently,

holding λ fixed at 1 assures that a change in achievable welfare is not driven by a change in the marginal cost of public funds.[21] As a consequence, as we vary x, we are also adjusting the share of high-skilled individuals in the population according to $\eta = \frac{1+x}{2}$.

i) For $0 \leq x \leq 0.236$, $X^* = \{(\theta_H, w_L), (\theta_H, w_H)\}$, and $EW^{**}_{X^*}$ is a decreasing function of x. The optimal provision rule has *four bunching regions and is increasing for medium levels of* p, i.e., there exist numbers Q_{ss}, Q_s, Q_l and Q_{ll} such that

$$Q(p) \quad = \quad \begin{cases} Q_{ss} & \text{for} \quad 0 \leq p \leq \hat{p}\ , \\ Q_s & \text{for} \quad \hat{p} < p < \hat{p}'\ , \\ Q^*(p) & \text{for} \quad \hat{p}' \leq p \leq \tilde{p}'\ , \\ Q_l & \text{for} \quad \tilde{p}' < p < \tilde{p}\ , \\ Q_{ll} & \text{for} \quad \tilde{p} \leq p \leq 1\ , \end{cases}$$

where $U(Q_{ss}, \theta_L, w_H) = U(Q_s, \theta_L, w_H)$, and $U(Q_l, \theta_H, w_L) = U(Q_{ll}, \theta_H, w_L)$.

ii) For $0.236 \leq x \leq 0.244$, $X^* = \{(\theta_H, w_L), (\theta_H, w_H)\}$, and $EW^{**}_{X^*}$ is a decreasing function of x. The optimal provision rule has *three bunching regions*, i.e. there exist numbers Q_s, Q_m and Q_l such that

$$Q(p) \quad = \quad \begin{cases} Q_s & \text{for} \quad 0 \leq p \leq \hat{p}\ , \\ Q_m & \text{for} \quad \hat{p} < p < \tilde{p}\ , \\ Q_l & \text{for} \quad \tilde{p} \leq p \leq 1\ , \end{cases}$$

where $U(Q_s, \theta_L, w_H) = U(Q_m, \theta_L, w_H)$ and $U(Q_m, \theta_H, w_L) = U(Q_l, \theta_H, w_L)$.

iii) For $0.244 \leq x \leq 0.373$, $X^* = \{(\theta_H, w_H)\}$, and $EW^{**}_{X^*}$ is an increasing function of x for $x < \frac{1}{3}$ and a decreasing function for $x > \frac{1}{3}$. The optimal provision rule has *two bunching regions and is increasing for high levels of* p, i.e., there exist numbers Q_s and Q_l such that

$$Q(p) \quad := \quad \begin{cases} Q_s & \text{for} \quad 0 \leq p \leq \hat{p}\ , \\ Q_l & \text{for} \quad \hat{p} < p < \hat{p}'\ , \\ Q^*(p) & \text{for} \quad \hat{p}' \leq p \leq 1\ . \end{cases}$$

If $\theta_L w_H \leq \theta_H w_L$, then $U(Q_s, \theta_H, w_L) = U(Q_l, \theta_H, w_L)$, and if $\theta_L w_H > \theta_H w_L$, then $U(Q_s, \theta_L, w_H) = U(Q_l, \theta_L, w_H)$.

iv) For $0.373 \leq x \leq 1$, $X^* = \{(\theta_H, w_H)\}$ and the optimal provision rule has *two bunching intervals* with provision levels Q_s and Q_l. For $0.373 \leq x \leq \frac{1}{2}$, the constraint $U(Q_s, \theta_L, w_H) \geq U(Q_l, \theta_L, w_H)$ is not binding so that $EW^{**}_{X^*}$ is constant. For $x \geq \frac{1}{2}$, the constraint $U(Q_s, \theta_L, w_H) = U(Q_l, \theta_L, w_H)$ is binding implying that $EW^{**}_{X^*}$ is a decreasing function of x. For $x \to 1$, $EW^{**}_{X^*}$ converges to the maximal level of welfare that can be achieved by choosing the same provision level for every p.

[21] However, none of our claims about comparative statics results relies on this simplification. It is made purely for the sake of convenience.