

# Why Are Some Taxes “more equal than others”?\*

Anna Rubinchik-Pessach<sup>†</sup>

Revised: June 2002

## Abstract

This paper examines efficiency rationales behind uniformity of taxation for public projects.

Let a benevolent risk neutral “constitutional designer” to set an optimal cost-sharing rule for a legislature operating under majority rule. Then the designer will choose ‘more equal taxes’ for a country with more homogeneous tastes, which is in accord with a popular view. Higher quality projects provide an additional reason for this choice. Moreover, an exogenous requirement to use broader supermajority may also lead to more uniformity.

**Keywords:** public goods, cost sharing, constitutional design

**JEL Classification Codes:** H41, H21, H61, K34, D72

---

\*I would like to thank Andrew Postlewaite for his guidance. This work was revised while the author was visiting CORE, the hospitality of which is greatly appreciated. It was partially supported by the European Commission Network - HP RN-CT-2000-00072. I have benefited from discussions with Jean Hindriks, George Mailath, Maurice Marchand, Steve Matthews, Nicola Persico, Amir Rabah and Shlomo Weber. Comments of the participants of the Applied Microeconomics Seminar at the University of Pennsylvania and the ASSA Meetings in New Orleans are appreciated. Any remaining errors are mine.

<sup>†</sup>University of Colorado at Boulder, Department of Economics, UCB 256, Boulder, CO, 80309-0256, USA. E-mail Anna.Rubinchik@Colorado.edu

## 1. Introduction

Equal cost sharing. Why is it so frequently used?

No doubt it is fundamental for a variety of budgeting institutions. In the US, for example, the receipts from income, corporate and excise taxes are classified as “on-budget”, thus not earmarked to be spend on a certain program or in a given region.<sup>1</sup> The tax code describing the rules of federal tax collection does not discriminate across regions either. In fact, Constitution of the United States explicitly forbids such a discrimination.<sup>2</sup> As for European countries, “common pool” of tax revenues lies at the core of budgeting process as well.<sup>3</sup>

This phenomenon is even more surprising in the view of the recent developments in the political economy literature. As acknowledged by many authors, equal cost sharing gives rise to misallocation of resources by a (central) government. See Baron (1991), Besley and Coate (2000), Boadway and Flatters (1982), Chari, Jones, and Marimon (1997), Ferejohn, Fiorina, and McKelvey (1987), Inman and Fitts (1990), Lockwood (2002) among others. In a stylized environment a government is comprised of democratically elected representatives who are collectively responsible for allocating public funds. In this setting the representatives are tempted to get central government financing for projects benefitting their constituents, often times without generating substantial benefits for the other residents of the country. Considering the project the politicians compare marginal benefit generated by their (local) project to just a fraction of marginal cost, instead of its full amount. By doing so they impose a cost on the other regions, who contribute to the general tax revenues. This creates a budgetary externality, which is a direct consequence of equal cost sharing. Why is this arrangement used then? Is there an efficiency rationale for it, is there a gain that leads a society to impose equal cost sharing?

Adopting the stylized framework, this paper starts with a model of a country governed by a legislature consisting of regional representatives, who make decisions based on majority voting. The role of the government is to choose public projects

---

<sup>1</sup>Social Security and the Postal Service are the only two programs excluded from the budget, in other words, are currently “off-budget”. “On-budget” receipts have constituted roughly 3/4 of the total receipts since 1975. “On-budget” spending has been around 80% of the total spending for the same time period. See Office of Management and Budget *Historical tables* publicly available at <http://w3.access.gpo.gov/usbudget>.

<sup>2</sup>Section 8,. Clause 1; Section 9, Clause 4; Amendments, Article XVI.

<sup>3</sup>See von Hagen (1992) for details and further references.

to provide and to finance them through taxation. The cost sharing rules are written in a “constitution”. The key difference introduced in this paper is that the tax burden does not have to be equal across regions. It may be (close to) uniform only if optimal, i.e., if it induces the government to make “better” decisions with respect to public goods provision.

It is natural in this respect to follow contractarian tradition to justify equal cost sharing as a “constitutionally” imposed arrangement. This approach became popular in rationalizing a wide spectrum of social arrangements starting from formally written laws (constitutions) to informal rules of sharing surplus from economic activities. (See Buchanan and Tullock (1962), Binmore (1998), Bös and Kolmar (2000) among others.) These arrangements are justified only if they are a part of a contract signed by the participants, in some “initial situation”, or behind the “veil of ignorance”. In other words, laws should be an epitome of common interest.

In accord with this tradition, let us assume that constitution creators are legislators behind the “veil of ignorance” (à la Rawls (1971)): they are yet unaware of their “personal position” (identified by the group of constituents they represent). One can follow Harsanyi (1992) in assuming that the individuals are rational and that it is common knowledge that each one of them may occupy any position (represent any locality with equal chance) after the “veil of ignorance” is lifted. Besides, they share common prior beliefs with respect to appearance of possible projects on the agenda and, moreover, the distribution describing these beliefs is symmetric. The benefits and the costs of a project are expressed in the same (monetary) terms, thus, by Harsanyi (1992), we can conclude that the social objective is to maximize the expected value of a possible project. As all the ex-ante identical individuals agree on their objective, we can view a ‘group of legislators behind the veil of ignorance as ‘constitutional designer’. The question then becomes: “What will motivate the designer to shrink possible gaps in cost shares for public projects?”

It is true that risk aversion of the legislators behind the veil of ignorance can lead them to agree to split all the costs equally for any project. To focus on additional explanations, risk neutrality is assumed.

At the time the “constitution” is written, the exact specification of public projects to be voted upon is not known. The only relevant information the designer has at that time is the probability distribution over possible projects. The source of the underlying uncertainty is twofold. First, “external” factors, such as earthquakes, climate changes, etc., can affect the frequency of the public projects

appearance. Second, “internal” factors, for example, budgetary rules and government organization, may shape the proposed projects. An aggregate impact of all these factors is taken as a “black box” in this model, being reduced to the definition of a probability distribution over projects.

It is the shape of this distribution that will be reflected in the constitutional rules. Clearly, if any project ever considered by the legislature is valued identically by all the citizens, equal cost sharing is ideal: apart from being “fair”, it motivates the politicians to accept only those projects that generate positive welfare and reject all the rest. This can be the case if all the spending programs are devoted to “pure public” goods and the willingness to pay for them is equal across regions. This echoes a well known result in fiscal federalism literature (Oates (1972), Besley and Coate (2000), Lockwood (2002)) stating that public goods with high inter-regional spillovers will be provided more efficiently by central government (using uniform cost sharing), as opposed to decentralized provision, under which only one region is responsible for covering the costs. Thus, the first reason for the taxes to be close to uniform is that identically valued projects are relatively frequent on the agenda.

The other two reasons are somewhat less straightforward. When projects generate higher benefits (relative to the costs), the optimal sharing rule should become more uniform. Moreover, the bigger the required size of the minimum winning coalition, the smaller is the optimal tax levied on a supporter of the project, which may also contribute to a more “uniform” taxation.

The rest of the paper is organized as follows. The next section contains the description of the model and the analysis of the legislative behavior under a given constitution. In particular, the relationship between the constitutional cost sharing and the set of projects accepted by the legislature is established there. This enables to formulate the problem of the designer in section 2.4 and to perform “comparative statics” aimed at illustrating the need for a more uniform cost sharing. Extensions offer an introspection into the problem of the designer; an environment, in which both majority voting and uniform taxation are optimal is illustrated there. Naturally, the last section concludes.

## 2. The Model

### 2.1. Environment

Let  $R \geq 3$  denote the number of regions,  $R$  being odd. Residents of each region have identical attitudes towards public goods, but the attitudes across regions differ.

Each region has one representative in the central government, later referred to as legislature. This governing body is authorized to impose taxes (according to pre-determined constitutional rules) and to allocate the proceeds to public projects accepted by a majority of legislators. A project is associated with the profile of benefits  $b = (b_1, \dots, b_R)$  across the regions. Some public goods can be equally valuable to all the regions, for example, national defense. On the other hand, some projects can generate “concentrated” benefits, so that only a handful of regions are ready to pay for them, while the rest do not value them at all. A local bridge or a river dam could be an example of such a project.

To ease the exposition, the benefits are normalized, so that all the projects have the same per-region cost, unity. Thus,  $b_i$  is a “benefit to cost ratio”, i.e., the ratio of the region  $i$ 's benefit over the per-region cost. Alternatively,  $b_i$  can be thought of as a return on a tax dollar per region.<sup>4</sup>

Denote by  $F$  an ex-ante probability distribution over the set of projects,  $\{b \in [\underline{B}, \bar{B}]^R\} \subset \mathbb{R}^R$ . This distribution is a primitive of the model: it describes environment, in which legislature makes decisions. It reflects frequency with which different projects appear on the agenda. The paper focuses on establishing the relationship between the result of this interplay ( $F$ ) and the optimal cost sharing rule.

Given a probability distribution  $F$  the designer defines acceptance criteria for public projects. The following assumptions impose constraints on the choice of the designer.

**Assumption VO** Project is accepted if and only if at least  $m = \frac{R+1}{2}$  legislators vote for it;

---

<sup>4</sup>It is common in the public debate to measure state (region) benefit per dollar of tax contributions as the ratio of federal spending by state per state tax dollar, see publications of “Tax foundation”, for example (at <http://www.taxfoundation.org>). As the amount of public *spending* may not be a proper measure of the *benefit*, this is *not* what is captured by  $b_i$ . It is assumed in this model that regional residents have a well defined willingness to pay for a public project, i.e., a reservation value for it,  $B_i$ . Then  $b_i = B_i / (C/R)$ , where  $C$  is the cost of the project.

**Assumption TAX** Taxes are imposed only if the project is accepted;

**Assumption AN** The payment scheme is anonymous, i.e., payment can not depend on the name of the region;

**Assumption BB** The payment scheme is budget balanced;

**Assumption SPM** The Supporters of the project *Pay More* or at least as much as those who oppose the project.

In the rest of the paper these assumptions (**VO**, **TAX**, **AN**, **BB**, **SPM**) will be referred to as “initial assumptions”, **A0**.

Assumption **VO** restricts the ability of legislators to express the degree of their support for the projects: they can either vote “yes” or “no”.<sup>5</sup> This, coupled with anonymity, translates into a severe restriction on the set of acceptance/rejection procedures that the designer is can choose from. Clearly, this assumption can be justified on positive grounds. Majority voting is commonly used in practice, it is considered to be simple, and it may be dictated by reasons not directly related to economic efficiency.<sup>6</sup>

Anonymity can be viewed as a desirable one restriction as well. It assures that if two different legislators vote in the same fashion, they will end up paying the same tax, so that they can not be discriminated based on their identity.<sup>7</sup> This implies that the designer can differentiate the payments *only* on the basis of voting decisions of the legislators, thus being free to set just two levels of taxes: that for the supporters and the one for those who oppose the project.

Budget balance along with the second assumption suggest that project’s costs have to be fully covered by the tax proceeds, so that no “third party” is needed to finance the project in case of a deficit or to “burn the money” in case of a surplus. The last assumption can be justified on a “fairness” grounds. In fact, there is a reason to impose this restriction on efficiency grounds as well: it prevents the legislature to accept all the projects that appear on the agenda. This will be demonstrated in the discussion that follows proposition 2.1.

---

<sup>5</sup>It is assumed that the legislators can use only one vote per issue, they can not store votes as in Casella (2002).

<sup>6</sup>If maximizing the pivotalness of each voter is the objective, it is the best rule, see Al-Najjar and Smorodinski (2002).

<sup>7</sup>This is true provided the payment scheme is deterministic. Restricting attention to deterministic procedures eliminates the necessity to discuss reliability and observability of randomization devices and simplifies the exposition.

The rest of this section is devoted to the description of the set of accepted projects under assumptions **A0**.

## 2.2. Specification of the sharing rule.

Recall that due to normalization of benefits, per-region cost for all the projects is unity, thus, the total cost for all the projects equal to the number of regions,  $R$ . By anonymity, taxes may differ only on the basis of a legislator's voting decision: 'yes' or 'no',  $\{Y, N\}$ . So if  $k \geq m$  legislators vote in favor of the project, the designer has to set two levels of taxes:  $t_k(Y)$  and  $t_k(N)$ . Incorporating the *BB* restriction,

$$kt_k(Y) + (R - k)t_k(N) = R,$$

he is left with setting just a cost sharing parameter between the supporters and the 'no' voters. Denote this parameter by  $\alpha$ . If the taxes are uniform, then the share of the supporters is  $k$ . If they have to bear a higher share, then they pay  $\alpha_k k$ , so that  $\alpha_k = 1$  corresponds to the case, when the taxes are uniform. By assumption *SPM*,  $\alpha_k \geq 1$ . So, if  $k \geq m$  people vote for the project, each one of the supporters pay

$$t_k(Y) = \alpha_k,$$

where  $\alpha_k \geq 1$  and  $\alpha_R = 1$ ; whereas anyone who opposed the project is required to pay

$$t_k(N) = \frac{R - \alpha_k k}{R - k}.$$

Thus, the sharing rule is fully specified by the vector  $\alpha = (\alpha_m, \alpha_{m+1}, \dots, \alpha_R)$ . The section demonstrates that we can restrict attention only to  $\alpha_m$  to determine the set of outcomes of the voting game, and thus this is the only parameter that will control the efficiency of the mechanism. In what follows, we will compare optimal tax schemes for different environments using  $\alpha_m$ . Based on our definition of this parameter, lower optimal  $\alpha_m$  corresponds to a more uniform taxation.

## 2.3. Voting Stage

Faced with the cost sharing arrangement,  $(\alpha_m, \alpha_{m+1}, \dots, \alpha_R)$ , and given a realization of a project,  $(b_1, \dots, b_R)$ , the legislators play a simultaneous full information voting game. In this game a legislator  $i$  has two actions (votes):

$a_i \in A = \{Y, N\}$ ; and a linear utility :

$$U_i(b_i, a_i, a_{-i}) = \left\{ \begin{array}{l} b_i - t_k(a_i), \text{ if } |\{j : a_j = Y\}| \geq m; \\ 0 \text{ otherwise.} \end{array} \right\}, \quad (2.1)$$

where  $|X|$  denotes cardinality of set  $X$ .

Recall that the project is accepted iff the number of legislators voting  $Y$  is at least  $m$ ,  $k \geq m$ . It is rejected otherwise.

Assume the legislators have full information about the benefits that a potential project generates (consider, for example, the existing legislative process in the US, for which this assumption is quite realistic).

Consider trembling hand perfect Nash equilibria of this game (Selten (1975)).

Denote by  $M$  the set of projects that will be accepted, if this equilibrium is being played. The following proposition demonstrates that this set can be fully described by the cost sharing parameter,  $\alpha_m$ . The description of this set is simple: the projects that generate  $m^{\text{th}}$  highest benefit above the threshold  $\alpha_m$  will be accepted by the legislature (whereas all the rest of the projects will be rejected). In particular, it implies that as long as  $\alpha_{m+1}, \dots, \alpha_R$  satisfy constraint *SPM* they are irrelevant for the description of the set of accepted projects.

**Proposition 2.1.** *Assume A0. Then the set of projects accepted under a trembling hand perfect Nash equilibrium can be described as follows:*

$$\{(b_1, \dots, b_R) : b_{[m]} > \alpha_m\} = M(\alpha_m). \quad (2.2)$$

The **proof** is relegated to appendix A.

The claim is based on the following result: whenever the project is provided, exactly  $m$  people vote for it provided the taxes are non-uniform, i.e., if the supporters pay strictly more than those who oppose the project. The same is true if the taxes are uniform for robust equilibria.<sup>8</sup>

It is worth noting that the simple majority assumption,  $m = (R + 1) / 2$ , was never used in the proof. Thus it is valid for any  $m$  between 1 and  $R$ .

Although condition *SPM* is crucial for characterization of the equilibrium outcomes, it is dispensable. Indeed, it implies that if  $k \geq m$  individuals vote for the proposal it is a dominant strategy for any other player to vote against, as this way he saves on his tax bill:  $t_{k+1}(Y) \geq t_k(N)$ . Relaxing *SPM* reverses this inequality (for some  $k \geq m$ ) and thus, gives rise to another set of equilibria under

---

<sup>8</sup>See the details in the appendix A.

which any project is accepted independent of its quality. Clearly, the equilibria in which the projects described by 2.2 are accepted will be retained as well. Provided accepting all projects is socially undesirable (generates negative expected value), the designer may want to refrain from violating *SPM* on efficiency grounds.

Trembling hand refinement eliminates “unreasonable” equilibria, in which nobody votes for the project (however good it is) or, more generally, the equilibria, in which independent of the benefits generated, only  $k < m$  legislators vote in favor of a project. I will leave it to the reader to recall anecdotal evidence about pushing a wrong (yes or no) button during voting. The fact that voters consider the possibility that others may err provides justification for using this refinement.

## 2.4. The Problem of the Designer

Recall that the benefit profile is normalized so that the per-region cost of a project is unity, thus, the utilitarian welfare from project  $b$  is  $\sum_{i=1}^R b_i - R$ . A project, generating positive welfare will be referred to as “efficient”. Slightly abusing notation,  $\alpha_m$ , the crucial component of the cost sharing vector will be denoted simply by  $\alpha \in \mathbb{R}$ .

The problem facing the designer can be written as follows,

$$\max_{\alpha \geq 1} \Phi(\alpha, m),$$

$$\Phi(\alpha, m) \equiv \Pr(b_{[m]} > \alpha) E\left(\sum_{i=1}^R b_i - R \mid b_{[m]} > \alpha\right)$$

Observe that cost sharing parameter  $\alpha$  controls the threshold for the  $m^{\text{th}}$  highest benefit, so that the higher is  $\alpha$ , the less projects (both efficient and inefficient) will be accepted by the legislature. Therefore, the designer will have to choose the sharing parameter in order to make just the right trade-off between the expected value of the excluded efficient and inefficient projects.

Another way to analyze this problem is to notice that increasing  $\alpha$  should improve the conditional expectation term, while decreasing the probability of a project to be accepted (the first term). If the threshold  $\alpha$  is “too high”, very good projects will be accepted, but a lot of projects will be rejected, including some “good” ones, thus inducing a considerable “type *II*” error. It may be worthwhile to decrease  $\alpha$ , thus extending the set of accepted projects at a cost of, sometimes, accepting “wrong” projects (“type *I*” error). On the other extreme, if  $\alpha$  is “too

low”, set of accepted projects being substantial, the severity of type  $I$  error may cause the designer to increase  $\alpha$ .

A reasonable restriction on the distribution function  $F$  is sufficient for existence and uniqueness of the solution to this problem. First, let us define affiliation as in Milgrom and Weber (1982).

**Definition 2.2.** *The random variables  $\{b_1, \dots, b_R\}$  are affiliated if the joint density  $f(b)$  is such that for any  $b, b'$*

$$f(b \wedge b')f(b \vee b') \geq f(b)f(b'),$$

where  $b \vee b'$  is component-wise maximum,

$$b \vee b' = \left( \max \{b_1, b'_1\}, \dots, \max \{b_R, b'_R\} \right)$$

and  $b \wedge b'$  is component-wise minimum,

$$b \wedge b' = \left( \min \{b_1, b'_1\}, \dots, \min \{b_R, b'_R\} \right).$$

Affiliation implies that if one of the components of the vector is high (low), it is more likely that the others are high (low) as well. Assuming that regional benefits are affiliated is consistent with the idea that a public project has an ‘objective’ (common) value perceived differently by regions. Variation in the willingness to pay for the project can be a result of differences in tastes, as well as technological constraints. The appeal of a publicly broadcasted ballet can vary across regions according to their tastes, whereas the benefits of environmental regulations can vary according to their exact (technological) specifications. Note that it is still possible to have projects with *very* different regional valuations. Affiliation just requires these occurrences to be less frequent. Moreover, it does not rule out independently distributed regional valuations.

**Assumptions A1** Assume the benefits are distributed on  $[\underline{B}, \bar{B}]^R$ ;  $\underline{B} < 0$  and  $\bar{B} > R > 0$ ; with the corresponding strictly positive continuous p.d.f. on  $[\underline{B}, \bar{B}]$ . Assume this distribution is symmetric. Assume also that  $\{b_1, \dots, b_R\}$  are affiliated.

Most of the following statements use the following theorem from Milgrom and Weber (1982).

**Theorem** <sup>9</sup> Let  $b_1, \dots, b_R$  be affiliated and let  $H$  be any non-decreasing function. Then the function  $h$  defined by

$$h(a_1, c_1; \dots, a_R, c_R) = E(H(b_1, b_2, \dots, b_R) | a_1 \leq b_1 \leq c_1, \dots, a_R \leq b_R \leq c_R)$$

is non-decreasing in all of its arguments.

**Proposition 2.3.** Assume the distribution over benefits  $F$  satisfies **A1**. Then a maximizer of  $\Phi(\alpha, m)$  exists and is unique.

**Proof.** First, note that it is never optimal to set  $\alpha > R/m$ . Indeed, in the view of proposition 2.1,  $\alpha = R/m$  implies that only projects generating positive ex-post utilitarian welfare will be accepted,

$$\begin{aligned} b_m &> R/m \Rightarrow \\ \sum_{i=1}^R b_i &\geq \sum_{i=1}^m b_i \geq mb_{[m]} > R. \end{aligned}$$

Therefore, increasing threshold  $\alpha$  above  $R/m$  will strictly decrease welfare by eliminating some desirable projects from the set of those accepted. Hence, without loss of generality we can require  $\alpha \leq R/m$ , thus, along with the restriction  $\alpha \geq 1$ , it amounts to  $\alpha \in [1, R/m]$  assuring existence of the solution in the view of continuity of the objective function  $\Phi(\alpha, m)$  in the first argument and the Weierstrass theorem.

Moreover, given **A1**, by lemma C.1, the first order conditions for the maximization problem can be represented as

$$\begin{aligned} \Phi_\alpha(\alpha, m) &\leq 0, \\ \Phi_\alpha(\alpha, m) &= -m \binom{R}{m} P(\alpha, m) V(\alpha, m) \end{aligned} \tag{2.3}$$

where

$$V(\alpha, m) \equiv E(w(b) + \alpha | \mathcal{A}(\alpha, m)) \tag{2.4}$$

$$w(b) = \sum_{i \neq m} b_i - R \tag{2.5}$$

$$P(\alpha, m) \equiv \Pr(b \in \mathcal{A}(\alpha, m)), \tag{2.6}$$

$$\mathcal{A}(\alpha, m) = \{b : b_1, \dots, b_{m-1} > \alpha, b_m = \alpha, b_{m+1}, \dots, b_R < \alpha\}. \tag{2.7}$$

---

<sup>9</sup>Theorem 5, in Milgrom and Weber (1982), p.1100

Clearly, if the constraint  $\alpha \geq 1$  is binding,  $\Phi_\alpha(\alpha, m) < 0$ ; if the solution is interior, then  $\Phi_\alpha(\alpha, m) = 0$ .

Moreover, by the lemma C.1,  $P(\alpha, m) > 0$  for  $\alpha \in [1, R/m]$  and  $V(\alpha, m)$  is strictly increasing in  $\alpha$  and, therefore can change sign at most once on the set of feasible values of  $\alpha$ . ■

Existence and uniqueness of the solution to the designer’s problem allows to perform meaningful “comparative statics” experiments on the primitive of the model, the probability distribution  $F$  defined over projects that appear on the agenda.

### 3. The Results

Recall that assumption *SPM* prescribes the supporters of the project to pay no less than those who vote against it. As stated before, it translates into a restriction of the set of feasible values of the sharing parameter,  $\alpha$ , to be above unity. Thus, a (weak) decrease in the optimal parameter will correspond to more uniform taxation. Indeed, as  $\alpha$  approaches 1, the tax levied on any supporter of the project,  $t_m(Y) = \alpha$ , will approach the tax levied on anyone who voted against it,

$$t_m(N) = \frac{R - \alpha m}{R - m}.$$

This allows us to formulate the results in terms of monotonic properties of the optimal parameter.

#### 3.1. Homogeneity of Benefits

As one would expect, more homogeneous tastes require more uniform taxes under majority voting. In the “extreme case”, if all the legislators receive the same benefit from the project they should pay the same tax in the optimum. The following proposition describes a monotonic relationship between the degree of homogeneity of benefits and the degree of tax uniformity.

To proceed, we need a suitable notion of homogeneity. It is clear that in this setting the degree of homogeneity of tastes is determined by a property of the distribution over the profile of benefits. Say, tastes are more homogeneous if there is a higher chance that all the members of the legislature value the project exactly the same, i.e., all the realizations of the profile are identical. Let’s use  $\lambda \in [0, 1]$  to

parametrize this factor. Here, when  $\lambda = 1$ , the benefits are purely homogeneous, and, as this parameter decreases, the benefits become more heterogeneous.

**Proposition 3.1.** *Assume **A0**. Consider a random variable  $v$  which is distributed  $G$  on  $[\underline{B}, \bar{B}]$ ;  $\underline{B} < 0, \bar{B} > 0$ . Assume that, conditional on  $v$ , the distribution of benefits is as follows: with probability  $\lambda$  benefits are identically equal to  $v$ , and with probability  $(1 - \lambda)$  they are distributed  $F$  satisfying **A1**.*

*Then as  $\lambda$  grows, the global maximizer of the objective function,  $\alpha$ , will decrease or stay constant.*

**Proof.** See appendix *D*.

The restriction that all the regional benefits are exactly equal to each other with positive probability may be too strict or hard to satisfy in practice. To show robustness of the statement, it is possible to formulate similar property requiring the benefits to be “very close” to each other with positive probability. It is possible to assure that the cost sharing parameter  $\alpha$  will decrease as this probability increases, but may not be monotonic as it approaches unity.<sup>10</sup>

### 3.2. Better Projects

This subsection is devoted to a somewhat less straightforward result, suggesting a possible rationale for uniformity of taxation under majority voting. The proposition 3.3 shows that if there is “more at stake”, the designer will set optimal taxes in a more uniform fashion. Indeed, when the expected project becomes better, it is optimal to increase the chances of it being accepted by lowering the taxes on the members of the minimum winning coalition. Note that the discrepancy in taxes create the “free rider” effect: the ‘yes’ voters have to pay more than the ‘no’ voters, in case the public good is provided. The discrepancy is needed to take care of the external costs that the majority imposes on the minority by obliging it to accept the project it deems worthwhile. So, if the projects become better, the “free rider” effect becomes more severe, while possible negative consequences of the imposed externality are smaller. In this case the designer wishes to reduce the differentiation of taxes, hence making them more uniform.

First, consider an example.

**Example 3.2.** *Assume the benefits are i.i.d. uniform on  $[\underline{B}, \bar{B}]$ ,  $\bar{B} > 3, \underline{B} \leq 0$ . Assume  $R = 3$  then  $m = (R + 1) / 2 = 2$ . Then  $\underline{B} < 1 \leq \alpha \leq 3 < \bar{B}$ . It can be*

---

<sup>10</sup>The author will gladly provide the details.

shown (see (B.2)) that the first order condition in this case can be represented as

$$\begin{aligned}\Phi_{\alpha}^{unif}(\alpha, 2) &= -K(\alpha - \underline{B})(\bar{B} - \alpha) \left( 2\alpha - 3 + \frac{1}{2}(\underline{B} + \bar{B}) \right) = 0 \Rightarrow \\ \Rightarrow \alpha_2^* &= \frac{3 - \frac{1}{2}(\underline{B} + \bar{B})}{2},\end{aligned}\tag{3.1}$$

where  $K$  is a positive constant. It is easy to see that the internal maximizer,  $\alpha_2^*$ , is decreasing in the expected benefit,  $\frac{1}{2}(\underline{B} + \bar{B})$ , so that the taxes become more uniform as the projects become better. Note that if the expected per-region benefit is sufficiently high,

$$\frac{1}{2}(\underline{B} + \bar{B}) \geq 1,\tag{3.2}$$

then the taxes should be uniform, i.e., the “corner solution” will be chosen.

Next, let us generalize the idea. Consider an environment, in which both “bad” and “good” projects can appear on the agenda. Compare it with another environment, in which no “bad” projects appear, instead, only “excellent” and “good” ones will be considered by the government. How should the optimal tax react to such a change? As example considered above suggests, the tax differential should shrink.

**Proposition 3.3.** *Assume A0. Consider two environments. In the first projects are distributed  $F$  on  $[\underline{B}, \bar{B}]^R$  and the second (with “better projects”) is characterized by distribution  $G$  on  $[\underline{B} + a, \bar{B} + a]^R$ , such that  $G(b) = F(b - a)$ . Assume  $G$  and  $F$  satisfy A1. Let  $\alpha^G, \alpha^H$  denote the optimal cost sharing rule under the distribution  $G, F$  respectively. Then  $\alpha^G \leq \alpha^H$ .*

**Proof.** See appendix *D*.

One could suggest that the idea of “improved” projects can be formulated in terms of first order stochastic dominance. Evidently, it is possible to construct a counter-example, demonstrating that there are first order stochastic dominance transformations of the original distribution over projects that lead to higher  $\alpha$  (more “differentiated” taxation), see appendix *E* for details.

The idea behind the counter-example is quite simple. The optimal tax is determined by the “marginal” condition stating that the welfare generated by the last accepted (rejected) project should be zero. Recall that tax differential  $\alpha$  is a threshold that determines the division of the set of all projects into accepted

and rejected. Under majority rule marginal project is the one that has its median benefit equal to  $\alpha$ , whereas half of the components above and the other half below this threshold. Fix an optimal threshold for a given distribution  $F$  and consider a first order stochastic transformation of  $F$  that shifts the probability mass from the projects just below the threshold (“not very bad”) to those just above (“moderately good”). Now, if a regional benefit falls below the threshold  $\alpha$ , it is really “bad news”, it is very likely that this benefit will have a very low value (“very bad”). Moreover, conditioning on the fact that the half of the benefits fall above the threshold is not as “good news” as before, as it is more likely that they fall “just” above the threshold. This decreases the expected welfare, thus, the value of the “marginal project” (which was zero before the transformation) becomes negative. This causes the designer to shift the threshold up, thus, making the taxes more differentiated, even though the projects became better on average.

### 3.3. Bigger Supermajority May Lead to a More Uniform Taxation

So far, we assumed that the decisions are being made by a simple majority of legislators. Requiring exactly half of the legislators to be decisive may not be always desirable. For example, there may be decisions that can have large adverse effects on some regions. In this case the designer may want to be sure that all (or almost all) of the legislators agree with the decision by setting supermajority or unanimity rule.

Let  $m$  now denote the size of the minimum winning coalition (which can be different from  $(R + 1)/2$ ). The following statement shows that the taxes may become more uniform as the required size of the minimum winning coalition grows.

**Proposition 3.4.** *Assume benefits are distributed  $F$  satisfying **A1**, also assume **A0** except **MJ**.*

*Then for any  $m \in \{1, 2, \dots, R - 1\}$*

$$\alpha_m \geq \alpha_{m+1}, \tag{3.3}$$

*where  $\alpha_m = \arg \max_{\alpha} \Phi(\alpha, m)$ ;  $\alpha_{m+1} = \arg \max_{\alpha} \Phi(\alpha, m + 1)$ .*

**Proof.** In the view of lemma C.1 and monotonicity of  $V(\alpha, m)$  in the first argument, it is sufficient to show that  $V(\alpha, m)$  is also monotonically increasing in the second argument,  $m$ . But given assumptions **A1**, this is a direct consequence of theorem 5 in Milgrom and Weber (1982). ■

In fact, this result is quite intuitive. Increasing the size of the minimum winning coalition or setting higher taxes on the coalition members will restrict the set of accepted projects, thus, increasing the value of wrongly rejected projects. To compensate for this, the optimal tax on the supporters should decrease allowing more projects to be accepted.

What does this result imply about uniformity of taxation? So far it was argued that a decrease in the cost sharing parameter  $\alpha_m$  corresponds to more uniform taxation. This is true for a fixed size of the minimum winning coalition,  $m$ . If the size of this coalition is not fixed, there is an additional requirement that has to be satisfied for the taxes to be “more uniform”. Note that inequality 3.3 just guarantees that supporters will pay smaller fraction of the average cost. As the number of supporters grow, it can have an ambiguous implication on the relative share paid by the ‘no’ voters. It can be easily shown that if

$$\alpha_{m+1} \leq \frac{\alpha_m (R - m)}{R - m + \alpha_m - 1}, \quad (3.4)$$

then the ratio of the taxes paid by supporters and the ‘no’ voters will decrease as a result. If this ratio is taken as a measure of tax uniformity, condition 3.4 implies that the optimal cost sharing parameter must decrease “enough” to make the taxes more uniform as a result.

## 4. Extensions

### 4.1. When is Uniform Tax Optimal?

Uniform taxation and simple majority rule can be a good approximation for the optimal mechanism for an interesting class of distributions over benefits as the population of voters (the number of the legislators in this case) is arbitrarily large. The distribution of the benefits can be interpreted as being generated by some “common value”,  $v$ , and the “region specific noise” distributed symmetrically around  $v$ .

**Proposition 4.1.** *Consider a real valued random variable  $v$  distributed with an arbitrary distribution  $G$ . Assume a regional benefit  $b_i = v + \varepsilon_i$ , where  $\varepsilon_i$  are distributed identically and independently, with a symmetric distribution around the zero mean, so that  $E(b_i|v) = v$ .*

*Then as the number of regions grow infinitely large,  $R \rightarrow \infty$ , it is optimal to set majority rule with uniform taxes.*

**Proof.** is in the appendix C.

This result is closely related to the Condorcet jury theorem. Notice that  $v > 1$  implies the project is worthwhile (for infinitely high  $R$ ), so that the “correct” decision is to accept it. In this case a legislator gets a “correct” realization  $b_i > 1$  with probability higher than a half. Moreover, for a fixed  $v$  the distribution of signals is independent. But this corresponds to the hypothesis of the Condorcet jury theorem. The same assertion is true for the case  $v < 1$ . Clearly, in the current framework  $b_i$  has a different interpretation, as  $b_i$  denotes the realized regional benefit, rather than a “signal” about the future benefit from a project.

#### 4.2. Uniform Tax under Unanimity and “Dictatorship”

In general there is nothing that should prevent the designer to choose any size of the minimum winning coalition, not necessarily equal to  $(R + 1)/2$ . He, can, for example consider a “dictatorship”,  $m = 1$ . It easy to show that if the size of the minimum winning coalition is unity, it is never optimal to set a uniform tax. The reason for that is that if  $\alpha = 1$  and  $m = 1$  the set of accepted projects,  $\{b \in \mathbb{R}^R, b_{[1]} > 1\}$  strictly includes the set of efficient projects,  $\{b \in \mathbb{R}^R, \sum_{i=1}^R b_i > R\}$ . The value of the marginal project,

$$P(1, 1) E \left( \sum_{i=2}^R b_i - R + 1 | b_2, \dots, b_R < 1 \right)$$

is strictly negative (implying that the first derivative of the objective function is strictly positive). In this case it always pays to increase the threshold  $\alpha$  to exclude more projects from the accepted set (possibly rejecting some efficient projects as a result).

On the other hand, unanimity with uniform taxes leads to acceptance of efficient projects only. By the same token, the marginal project in this case has a positive value. Thus, imposing uniformity under unanimity is ‘constraint efficient’ in this case.<sup>11</sup>

These observations are illustrated in figure 4.1 constructed for  $R = 2$ . It demonstrates that the set of accepted projects under unanimity and uniform taxation is a subset of efficient projects, while the set of projects accepted under  $m = 1$  and uniform taxes includes the set of efficient projects. Recall that the sets of

---

<sup>11</sup>Recall that the notion of efficiency used in this model is a utilitarian one. Clearly, the set of accepted projects under unanimity with uniform taxes exactly coincides with the set of ex-post Pareto efficient projects (if no transfers/side payments are allowed).

accepted projects, vary both by the size of the minimum winning coalition (the first argument) and by the cost sharing parameter (the second argument), which is set to unity in this example.

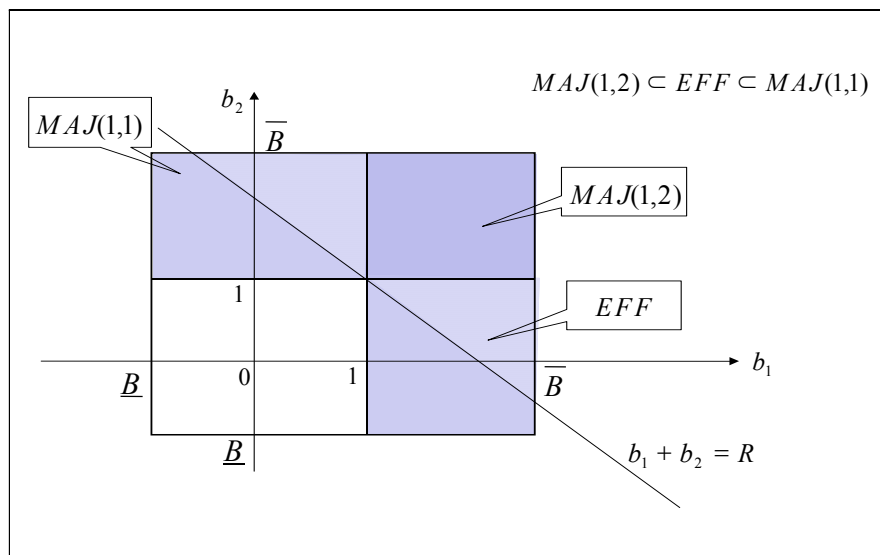


Figure 4.1: Relationship between the set of ex-post efficient projects (denoted by  $EFF$ ) and the sets of accepted projects (denoted by  $MAJ$ ) under two alternative acceptance rules.

### 4.3. Conclusions

Equal cost sharing for public projects may have a variety of explanations. This paper is focused on two of them. Homogeneity, or similarity of attitudes towards public projects is one and high quality of the projects, or the risk of rejecting good projects is another. These two reasons are valid if the acceptance/rejection decisions are based on voting with a fixed size of the minimum winning coalition. Evidently, requiring bigger coalitions to be decisive should be accompanied by reducing the share of costs imposed on the members of this coalition. This, in addition, can provide another reason for uniformity of taxation.

Recall that the results are stated for normalized benefits (benefits divided by per-region costs,  $b_i = B_i/c$ , where  $c = C/R$  and  $C$  is the total cost of the project.). They can easily be reformulated in terms of “pure benefits”  $(B_i)_{i=1}^R$

for any (ex-post) given cost  $C$  of a project. Moreover, if projects costs  $C$  and benefits  $(B_1, B_2, \dots, B_R)$  have common distribution  $\tilde{F}$  known to the designer, such that  $\{B_1, B_2, \dots, B_R, -C\}$  are affiliated, the results can be trivially reproduced in terms of expected benefits and costs, if so desired. Thus, normalization is purely for expositional purposes. This in contrast to other assumptions that appear to be crucial.

As suggested in the introduction, the explanation of the prevalence of equal cost sharing offered in this paper is based on contractarianism. A crucial component of this approach is specification of constraints imposed on the set of feasible contracts behind the veil of ignorance (see Bös and Kolmar (2000) for the relevant discussion). Recall that in the model regional benefits are commonly known by the legislators. It may not be that far from reality: often times the regional representatives get official estimates of the benefits that are expected to be generated by a project. But as Laffont and Maskin (1982) show under full information it is possible to construct optimal mechanisms for public good provision. Payment scheme will be (generally) uniform only if all the benefits are equal. Therefore, it is impossible to provide additional explanations for equal cost sharing (apart from homogeneity) in the “unrestricted world”. Therefore some constraints should be imposed to generate (almost) equal cost sharing as a commonly desirable rule.

The contribution of this paper is in establishing the connection between the set of constraints on constitutions, **A0** and the uniformity of taxes. Cornerstone assumptions in **A0** are voting and anonymity. Defending them lies beyond the scope of this paper, however.<sup>12</sup> Hence, provided there is a rationale for majority voting and anonymity, it is possible to justify uniform taxation.

One may wonder how crucial is the assumption about exogeneity of  $F$ . Certainly, as mentioned before, can, in part be viewed as a good positive feature of the model. Indeed, there are (external) factors that are beyond control of the legislators, the factors that can force them to consider public programs irrespective of the constitutional rules. It is the other group of factors that is worrisome in this respect. Having full understanding of the outcomes of the voting stage, legislators may have a motivation to change the formulation of the bills (public projects). That can definitely have an effect on the frequency with which projects of various kinds appear on the agenda. Packaging local projects (whether explicitly or through log-rolling) can be one of the examples of this phenomenon (see, Ferejohn, Fiorina, and McKelvey (1987), Inman and Rubinfeld (1997)). Moreover, as the

---

<sup>12</sup>Recent contributions in the area include Ledyard and Palfrey (2002), Al-Najjar and Smorodinski (2002).

distribution is “fixed”, appearance of a project on the agenda is independent on the decision made with respect to its predecessors.

This, however, does not eliminate the fact that the approach offered in the paper allows to identify important determinants of equal cost sharing embedded in the distribution over public projects. Whenever these determinants are present, they should be reflected in the optimal cost sharing. This relationship is valid, even though the distribution  $F$  may change over time.

Another question is whether equal cost sharing should be a persistent or a “stable” phenomenon. In other words, will the distribution  $F$ , the shape of which can vary with time, pertain necessary features that will lead to a (more) uniform taxation? Or, taking into account the connection between the constitution and the distribution of potential projects, will the legislators behind the veil of ignorance ever deem equal cost sharing worthwhile? The answers to these questions lie beyond the scope of the current paper, providing an avenue for future research.

## References

- Al-Najjar, N. I. and R. Smorodinski (2002). Pivotal Players and the Characterization of Influence. *Journal of Economic Theory*. forthcoming.
- Baron, D. P. (1991). Majoritarian Incentives, Pork Barrel Programs and Procedural Control. *American Journal of Political Science* 35, 57–90.
- Besley, T. and S. Coate (2000, May). Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Analysis. <http://econ.lse.ac.uk/staff/tbesley>.
- Binmore, K. (1998). *Game Theory and the Social Contract*, Volume 2: Just Playing. MIT.
- Boadway, R. I. and F. R. Flatters (1982). Efficiency and equalization of payments in a federal system of government: A synthesis and extension of recent results. *Canadian Journal of Economics* 15, 613–633.
- Bös, D. and M. Kolmar (2000, November). Anarchy, Efficiency and Redistribution. CESifo working paper no. 357.
- Buchanan, J. M. and G. Tullock (1962). *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: The University of Michigan Press.
- Casella, A. (2002). Storable votes. Columbia University, mimeo.

- Chari, V., L. E. Jones, and R. Marimon (1997). The Economics of Split-Ticket Voting in Representative Democracies. *The American Economic Review* 87, 957–976.
- Ferejohn, J., M. Fiorina, and R. D. McKelvey (1987). Sophisticated voting and agenda independence in the distributive politics setting. *American Journal of Political Science* 31, 169–194.
- Harsanyi, J. (1992). Game and decision theoretic models in ethics. In R. J. Aumann and S. Hart (Eds.), *Handbook of Game Theory With Economic Applications*, Volume 1. Amsterdam ; New York: North Holland.
- Inman, R. and M. Fitts (1990). Political Institutions and Fiscal Policy: Evidence from the U.S. Historical Record. *Journal of Law, Economics and Organization* 6, 79–132.
- Inman, R. P. and D. L. Rubinfeld (1997). The political economy of federalism. In D. C. Mueller (Ed.), *Perspectives on Public Choice: a handbook*. Cambridge: Cambridge University Press.
- Laffont, J. J. and E. Maskin (1982). The theory of incentives: An overview. In W. Hildenbrand (Ed.), *Advances in economic theory : invited papers for the Fourth World Congress of the Econometric Society at Aix-en-Provence*, pp. 31–94. Cambridge [Cambridgeshire], New York: Cambridge University Press.
- Ledyard, J. O. and T. R. Palfrey (2002). The Approximation of Efficient Public Good Mechanisms by Simple Voting Schemes. *Journal of Public Economics*. forthcoming.
- Lockwood, B. (2002). Distributive Politics and the Costs of Centralisation. *Review of Economic Studies* forthcoming.
- Milgrom, P. and R. Weber (1982). A Theory of Auctions and Competitive Bidding. *Econometrica* 50, 1089–1122.
- Oates, W. E. (1972). *Fiscal Federalism*. New York: Harcourt, Brace, Jovanovich.
- Rawls, J. (1971). *A Theory of Justice*. The Belknap Press of Harvard University Press.
- Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4, 25–55.
- von Hagen, J. (1992). Budgeting Procedures and Fiscal Performance in the European Communities. *Economic Papers* 96.

## A. Appendix

### A.1. Supporting the Outcomes of the Voting Game

**Proposition 2.1.** Assume **A0**. Then, under THPNE the set of the accepted projects has to satisfy:

$$M(\alpha_m) = \{b_{[m]} > \alpha_m\}.$$

**Proof.** We will start with a simplifying assumption that will be subsequently relaxed. Assume a stronger version of **(SPM)**, so that  $\alpha_k > 1$  for all  $k \geq m$ .

**Remark 1.** *There is no Nash equilibrium in which strictly more than  $m$  members vote for the project. Indeed, suppose  $R > k \geq m$  individuals vote for the proposal. By assumption  $\alpha_k > 1$ , so that  $\alpha_{k+1}(R - k) > R - k > R - \alpha_k k$ , thus  $t_{k+1}(Y) > t_k(N)$ . It follows that it is strictly dominant strategy for any player to vote against the project (to save on his tax bill).*

Define the following subsets of the set of legislators,  $\{1, \dots, R\}$  for a given project  $b$ :

- $LOW \equiv \{b_i < \alpha_m\}$  is the set of individuals with low valuation;
- $HIGH \equiv \{b_i > \alpha_m\}$  is the set of individuals with high valuation;
- $YES \equiv \{a_i = Y\}$  is the set of individuals voting for the project;
- $NO \equiv \{a_i = N\}$  is the set of individuals voting against the project.

Recall that the game  $G^b$  depends on the project under consideration,  $b$ . Thus, equilibrium of  $G^b$  may also depend on the project. The nature of this dependence is quite simple.

First, consider the case where the project is such that the  $m^{th}$  highest benefit it generates is relatively small, i.e.,  $b_{[m]} < \alpha_m$ . The following is the set equilibria in this case: not more than  $m - 1$  people are voting for the project while the rest oppose it. It's easy to see that there is no Nash equilibria in which  $|YES| = m$ , as in this case there must be at least one individual in the set  $YES$  who can profitably deviate by opposing the project, thus getting zero instead of a strictly negative payoff. According to our remark above, there is no NE such that  $|YES| > m$ . Therefore in any THPNE  $|YES| < m$ . So in this case the project is rejected.

Second, consider the case where the project is such that the  $m^{\text{th}}$  highest benefit it generates is relatively high, i.e.,  $b_{[m]} > \alpha_m$ . Notice that in this case a set of strategies where all those who support the project have high valuations (i.e.,  $YES \subset HIGH$ ) and  $|YES| = m$  is a THPNE. Indeed, it is a strict Nash equilibrium.

We are left to show that there is no THPNE in this case such that  $|YES| < m$ . It is obvious that any strategy profile supporting  $|YES| = m - 1$  is not a NE: there will be at least one person with high valuation who can profitably switch from opposing the project to supporting it, as  $|HIGH| > m$  in this case. Assume there is an equilibrium  $\sigma$  of the game such that  $|YES| = m - k$ ,  $k > 1$ . Consider an arbitrary sequence of purely mixed strategies  $\{\sigma^h\}_{h=0}^\infty$  that converges to  $\sigma$ . Let us show that for a person with high valuation,  $i \in HIGH$ , the choice of  $a_i = N$  is not a best response to  $\sigma_{-i}^h$  starting from some  $h$ . Indeed, if he chooses  $a_i = N$ , then it will take  $k$  trembles to approve the project, so that with probability of order  $(\varepsilon_h)^k$  he will get  $b_i - t_m(N)$ , while if less than  $k$  mistakes are made, he will get zero. On the other hand, if he switches to  $Y$ , then it will take only  $k - 1$  trembles to accept the project, giving him positive net benefit  $b_i - t_m(Y)$  with probability of order  $(\varepsilon_h)^{k-1}$ . Even if  $t_m(N) < t_m(Y)$ , as the difference between the taxes is bounded, when  $h$  is big enough ( $\varepsilon_h$  is small enough),  $a_i = N$  is not a best response to  $\sigma_{-i}^h$ .

The case  $b_{[m]} = \alpha_m$  will be of a little interest as I consider non-atomic distributions over the set of projects.

Thus we can conclude that

$$M(\alpha_m) = \{y \in Y : b_{[m]} > \alpha_m\}.$$

Now we can relax the assumption made at the beginning of the proof by allowing weak inequality  $\alpha_k \geq 1$ . Note that the only time we used the assumption about the strictness of inequality was to assure that  $t_{k+1}(Y) > t_k(N)$ . The only way to get equality  $t_{k+1}(Y) = t_k(N)$  is to have a vector of sharing rules,  $\alpha^* = (\alpha_m, \alpha_{m+1}, \dots, \alpha_{k^*}, \alpha_{k^*+1}, \dots, \alpha_{R-1})$ , such that  $\alpha_{k^*+1} = \alpha_{k^*} = 1$  for some  $k^* \geq m$ . Then, there will be equilibria in which  $k^* + 1$  legislators vote for the project given there are at least  $k^* + 1$  of them have valuations above the average cost. In addition, the equilibria of the second type, in which exactly  $m$  people vote for the project if  $b_{[m]} > \alpha_m$ , still remain. Note that if we take a sequence of vectors  $\alpha$  that converge to  $\alpha^*$  from above, then, only the latter equilibria will be present, while former will not, as along the sequence  $\alpha_k$  and  $\alpha_{k+1}$  are above 1 with at least one of them strictly above 1. So, the equilibria of the first type are not robust to the

“trembling hand of the designer”. Thus we are left with the equilibria supporting  $M(\alpha_m)$  as the set of accepted projects. ■

## B. Appendix: Existence and Uniqueness of the Solution

In the view of the assumptions **A1**, the objective function for a given  $R$ ,  $m \geq 1$ , can be represented as follows:

$$\Phi(\alpha, m) \equiv \sum_{k=m}^R \binom{R}{k} \underbrace{\int_{\alpha}^{\bar{B}} \dots \int_{\alpha}^{\bar{B}}}_{k \text{ times}} \underbrace{\int_{\underline{B}}^{\alpha} \dots \int_{\underline{B}}^{\alpha}}_{R-k \text{ times}} \left( \sum_{i=1}^R b_i - R \right) dF(b) \quad (\text{B.1})$$

Clearly, function  $\Phi(\alpha, m)$  is differentiable with respect to the first argument.

**Lemma B.1.** *Assume **A0**, **A1**. Then the first derivative of the objective function with respect to  $\alpha$  for fixed  $R, m$  can be represented as*

$$\Phi_{\alpha}(\alpha, m) = -m \binom{R}{m} A_1(\alpha, m), \quad (\text{B.2})$$

where

$$A_1(\alpha, m) \equiv \underbrace{\int_{\alpha}^{\bar{B}} \dots \int_{\alpha}^{\bar{B}}}_{m-1 \text{ times}} \underbrace{\int_{\underline{B}}^{\alpha} \dots \int_{\underline{B}}^{\alpha}}_{R-m \text{ times}} w(b, \alpha, m) dF(\alpha, b_{-m}) \quad (\text{B.3})$$

$$w(b, \alpha, m) = \sum_{i \neq m} b_i + \alpha - R, \quad (\text{B.4})$$

$$b_{-m} = (b_1, b_2, \dots, b_{m-1}, b_{m+1}, \dots, b_R). \quad (\text{B.5})$$

**Proof.** Taking first derivative of  $\Phi(\alpha, m)$  with respect to  $\alpha$  and using symmetry of the distribution  $F$ ,

$$\Phi_{\alpha}(\alpha, m) = \sum_{k=m}^R \binom{R}{k} \{-k * A_1(\alpha, k) + (R - k) * A_2(\alpha, k)\},$$

where

$$A_1(\alpha, k) \equiv \underbrace{\int_{\alpha}^{\bar{B}} \dots \int_{\alpha}^{\bar{B}}}_{k-1 \text{ times}} \underbrace{\int_{\underline{B}}^{\alpha} \dots \int_{\underline{B}}^{\alpha}}_{R-k \text{ times}} \left( \sum_{i \neq k} b_i + \alpha - R \right) dF(\alpha, b_{-k});$$

$$A_2(\alpha, k) \equiv \underbrace{\int_{\alpha}^{\bar{B}} \dots \int_{\alpha}^{\bar{B}}}_{k \text{ times}} \underbrace{\int_{\underline{B}}^{\alpha} \dots \int_{\underline{B}}^{\alpha}}_{R-k-1 \text{ times}} \left( \sum_{i \neq k+1} b_i + \alpha - R \right) dF(\alpha, b_{-(k+1)}).$$

First, note that

$$A_1(\alpha, k) = A_2(\alpha, k+1),$$

besides

$$\binom{R}{k} (R-k) = \frac{R!}{k!(R-k-1)!} = \binom{R}{k+1} (k+1).$$

Then for all  $k$  such that  $m \leq k \leq R-1$ ,

$$(R-k) \binom{R}{k} * A_1(\alpha, k) - (k+1) \binom{R}{k+1} * A_2(\alpha, k+1) = 0.$$

Moreover,

$$A_2(\alpha, R) = 0.$$

Therefore, after the cancellation, we get

$$\Phi_{\alpha}(\alpha, m) = -m \binom{R}{m} A_1(\alpha, m).$$

■

## C. Appendix: Auxiliary Statements

**Lemma C.1.** *Assume A1.*

*Then the first derivative of the objective function  $\Phi(\alpha, m)$  can be represented as*

$$\Phi_{\alpha}(\alpha, m) = -m \binom{R}{m} P(\alpha, m) V(\alpha, m), \quad (\text{C.1})$$

*where  $P(\alpha, m) > 0$  and  $V(\alpha, m)$  is strictly increasing in  $\alpha$  for a fixed  $m$ .*

**Proof.** Let

$$V(\alpha, m) \equiv \frac{1}{P(\alpha, m)} \underbrace{\int_{\alpha}^{\bar{B}} \dots \int_{\alpha}^{\bar{B}}}_{m-1 \text{ times}} \underbrace{\int_{\underline{B}}^{\alpha} \dots \int_{\underline{B}}^{\alpha}}_{R-m \text{ times}} [w(b) + \alpha] dF(\alpha, b_{-m}) \quad (\text{C.2})$$

$$w(b) = \sum_{i \neq m} b_i - R \quad (\text{C.3})$$

$$P(\alpha, m) \equiv \underbrace{\int_{\alpha}^{\bar{B}} \dots \int_{\alpha}^{\bar{B}}}_{m-1 \text{ times}} \underbrace{\int_{\underline{B}}^{\alpha} \dots \int_{\underline{B}}^{\alpha}}_{R-m \text{ times}} dF(\alpha, b_{-m}), \quad (\text{C.4})$$

$$b_{-m} = (b_1, b_2, \dots, b_{m-1}, b_{m+1}, \dots, b_R). \quad (\text{C.5})$$

Clearly, then

$$P(\alpha, m) V(\alpha, m) = A_1(\alpha, m),$$

where  $A_1(\alpha, m)$  is as defined in B.3, which justifies C.1. It is obvious that  $P(\alpha, m) > 0$  for all feasible  $\alpha \in [1, R/m] \subset [\underline{B}, \bar{B}]$ .

Therefore, it is left to show that  $V(\alpha, m)$  is strictly increasing in the first argument.

In the view of definition C.2,

$$V(\alpha, m) = E(w(b) + \alpha | \alpha < b_1, \dots, \alpha < b_{m-1}, b_{m+1} < \alpha, \dots, b_R < \alpha).$$

Pick  $\varepsilon > 0$ . Note that

$$V(\alpha + \varepsilon, m) > E(w(b) + \alpha | \alpha + \varepsilon < b_1, \dots, \alpha + \varepsilon < b_{m-1}, b_{m+1} < \alpha + \varepsilon, \dots, b_R < \alpha + \varepsilon).$$

Let

$$h(t) = E(w(b) + \alpha | t < b_1, \dots, t < b_{m-1}, b_{m+1} < t, \dots, b_R < t)$$

Then

$$V(\alpha + \varepsilon, m) - V(\alpha, m) > h(t + \varepsilon) - h(t).$$

Recall that by **A1**, random variables  $\{b_1, \dots, b_R\}$  are affiliated, besides,  $w(b)$  is increasing in  $b$ . Then, by Milgrom and Weber (1982), theorem 5,  $h$  is non-decreasing in  $t$ , which implies

$$h(t + \varepsilon) - h(t) \geq 0,$$

thus,  $V(\alpha, m)$  is strictly increasing in the first argument. ■

**Lemma C.2.** <sup>13</sup>Assume  $\Upsilon_1(\alpha) \geq \Upsilon_2(\alpha)$ ;

$$\begin{aligned}\alpha_1 &\equiv \arg \max_{\alpha} \int_{\alpha}^{\bar{B}} \Upsilon_1(t) dt; \\ \alpha_2 &\equiv \arg \max_{\alpha} \int_{\alpha}^{\bar{B}} \Upsilon_2(t) dt,\end{aligned}$$

Then

$$\alpha_1 \leq \alpha_2.$$

**Proof.** For any  $\beta < \alpha_1$ ,

$$\begin{aligned}\int_{\beta}^{\alpha_1} \Upsilon_1(t) dt &< 0 \Rightarrow \\ &\Rightarrow \int_{\beta}^{\alpha_1} \Upsilon_2(t) dt < 0 \Rightarrow \alpha_1 \leq \alpha_2.\end{aligned}$$

■

**Lemma C.3.** Consider the following equation

$$\Upsilon(\alpha, \lambda) = 0, \tag{C.6}$$

where

$$\Upsilon(\alpha, \lambda) \equiv \lambda p_1(\alpha) h_1(\alpha) + (1 - \lambda) p_0(\alpha) h_0(\alpha)$$

$\lambda \in [0, 1]$ ;  $\alpha \in [a, b] \subset \mathbb{R}$ ;

$p_i$  and  $h_i$  are continuous functions;  $p_i(\alpha) \geq 0 \forall \alpha \in [a, b]$ ,  $i = 0, 1$ ;  $h_i(\alpha)$  is strictly increasing in  $\alpha$ ;

One of the following conditions hold:

1.  $(\exists \alpha_1 \in [a, b] : h_1(\alpha_1) = 0 \text{ and } \exists \alpha_2 \in [a, b] : h_0(\alpha) = 0) \Rightarrow \alpha_1 < \alpha_2$ ;
2.  $\exists \alpha_1 \in [a, b] : h_1(\alpha_1) = 0 \text{ and } h_0(\alpha) < 0 \text{ for all } \alpha \in [a, b] \text{ (in this case let } \alpha_0 = b)$ ;
3.  $\exists \alpha_2 \in [a, b] : h_0(\alpha) = 0 \text{ and } h_1(\alpha) > 0 \text{ for all } \alpha \in [a, b] \text{ (in this case let } \alpha_1 = a)$ ;

---

<sup>13</sup>suggested by Nicola Persico

4. for all  $\alpha \in [a, b]$   $h_1(\alpha) > 0$  and  $h_0(\alpha) < 0$  (in this case let  $\alpha_1 = a$  and  $\alpha_0 = b$ )

Then

1. for any  $\alpha \in [\alpha_1, \alpha_0]$  either there exists a unique  $\lambda(\alpha)$  that solves equation (C.6).
2. for any  $\lambda_1, \lambda_0 \in (0, 1) : \lambda_1 > \lambda_0$  for which the corresponding sets of solutions  $A_{\lambda_1}, A_{\lambda_0} \subset [a, b]$  of the equation (C.6) are non-empty,  $\Upsilon(\alpha, \lambda_1) \geq \Upsilon(\alpha, \lambda_0)$ .

**Proof.**

Let's start with the first statement.

It follows from monotonicity and continuity assumptions and from definitions of  $\alpha_1, \alpha_0$  that for any  $\alpha \in [\alpha_1, \alpha_0]$   $h_1(\alpha) \geq 0$  and  $h_0(\alpha) \leq 0$ . Then

$$\lambda(\alpha) \equiv \frac{p_0(\alpha) h_0(\alpha)}{p_0(\alpha) h_0(\alpha) - p_1(\alpha) h_1(\alpha)}$$

solves (C.6) and is well defined in the sense that  $\lambda(\alpha) \in [0, 1]$ . It follows that  $\lambda(\cdot)$  is a continuous function of  $\alpha$ .

In the first case specified in the assumptions,  $\lambda(\alpha_1) = 1$  and  $\lambda(\alpha_0) = 0$ . Thus it follows that for any  $\lambda \in [0, 1]$  there exists at least one  $\alpha \in [\alpha_1, \alpha_0]$  that solves equation (C.6).

In the second case  $\lambda(\alpha_1) = 1$  and  $\lambda(\alpha) \geq \bar{\lambda} > 0$  for all  $\alpha$ . Thus for any  $\lambda \in [\bar{\lambda}, 1]$  there exists at least one  $\alpha \in [\alpha_1, \alpha_0]$  that solves equation (C.6) and for any  $\lambda \in [0, \bar{\lambda})$   $\Upsilon(\alpha, \lambda) < 0$ .

In the third case  $\lambda(\alpha_0) = 0$  and  $\lambda(\alpha) \leq \underline{\lambda} < 1$  for all  $\alpha$ . Thus for any  $\lambda \in [0, \underline{\lambda}]$  there exists at least one  $\alpha \in [\alpha_1, \alpha_0]$  that solves equation (C.6) and for any  $\lambda \in (\underline{\lambda}, 1]$   $\Upsilon(\alpha, \lambda) > 0$ .

In the last case  $0 \leq \underline{\lambda} \leq \lambda(\alpha) \leq \bar{\lambda} \leq 1$ , so that for any  $\lambda \in [\underline{\lambda}, \bar{\lambda}]$  there exists at least one  $\alpha \in [\alpha_1, \alpha_0]$  that solves equation (C.6), for any  $\lambda \in (\underline{\lambda}, 1]$   $\Upsilon(\alpha, \lambda) > 0$  and for any  $\lambda \in [0, \bar{\lambda})$   $\Upsilon(\alpha, \lambda) < 0$ .

Let's proceed with the second statement.

Assume that  $\lambda_1 > \lambda_0$  and  $\lambda_1, \lambda_0$  are in the relevant range, so that corresponding sets of solutions  $A_{\lambda_1}, A_{\lambda_0}$  for equation (C.6) are non-empty.

Then for any  $\alpha$

$$\Upsilon(\alpha, \lambda_1) \geq \Upsilon(\alpha, \lambda_0).$$

Indeed,

$$p_1(\alpha) h_1(\alpha) (\lambda_1 - \lambda_2) + p_0(\alpha) h_0(\alpha) (1 - \lambda_1 - 1 + \lambda_2) \geq 0.$$

■

**Proof.** of proposition 4.1. The designer has to find a subset of projects,  $Y^* \subset Y$  that maximizes the expected social welfare,

$$\begin{aligned} & \max_{Y^* \subset Y} E \left( \sum_{r=1}^R b_r - R \right) = \\ & = \max_{Y^* \subset Y} E_G \left[ E_H \left( \sum_{r=1}^R b_r - R | v \right) \right] = \\ & = \max_{Y^* \subset Y} E_G [Rv - R] = \max_{Y^* \subset Y} \int (Rv - R) dG(v) = \\ & = \max_{Y^* \subset Y} \left\{ \underbrace{\int_{v \leq 1} (Rv - R) dG(v)}_{\text{negative}} + \underbrace{\int_{v > 1} (Rv - R) dG(v)}_{\text{positive}} \right\}. \end{aligned}$$

To do so, the designer has to aim at getting only positive part of the sum. By Glivenko - Cantelli Law of Large Numbers,  $b_{[m]} \rightarrow v$ , as  $R \rightarrow \infty$ , i.e., the median ( $m = \frac{R+1}{2}$ ) of a realization is a good approximation for the median of the original distribution. Given that  $R$  is sufficiently large, the optimal set can be described as

$$Y^* = \{(b_1, b_2, \dots, b_R) : b_{(R+1)/2} \geq 1\},$$

which is precisely the set of projects accepted under simple majority rule with uniform taxes. ■

## D. Appendix: Proofs of the Main Results

**Proof.** of proposition 3.1.

First, let's write the objective function for this case, which we'll denote  $\Phi^\lambda$ .

$$\Phi^\lambda(\alpha, m) = \lambda \Phi^1(\alpha, m) + (1 - \lambda) \Phi^0(\alpha, m),$$

where

$$\Phi^1(\alpha, m) \equiv \int_{\alpha}^{\bar{B}} (Rv - R) dG(v)$$

and  $\Phi^0(\alpha, m)$  is as defined in (B.1).

Differentiating with respect to  $\alpha$ , we get

$$\Phi_{\alpha}^{\lambda}(\alpha, R) = \lambda \Phi_{\alpha}^1(\alpha, R) + (1 - \lambda) \Phi_{\alpha}^0(\alpha, R), \quad (\text{D.1})$$

where

$$\Phi_{\alpha}^1(\alpha, R) = -(\alpha R - R) G'(\alpha) \leq 0 \text{ for } \alpha \geq 1 \quad (\text{D.2})$$

and  $\Phi_{\alpha}^0(\alpha, R)$  is as defined in (B.2, B.3).

It is straightforward that the maximizer of  $\Phi^1(\alpha, R)$  is  $\alpha^1 = 1$ .

Maximizer of  $\Phi^0(\alpha, R)$  exists and is unique by proposition 2.3.

If the maximizer of  $\Phi^0(\alpha, R)$  is  $\alpha^0 = 1$ , then from (D.1), the maximizer,  $\alpha^{\lambda}$ , of  $\Phi^{\lambda}(\alpha, R)$  is also 1 for any  $\lambda \in [0, 1]$ .

If the maximizer of  $\Phi^0(\alpha, R)$ ,  $\alpha^0 > 1$ , then, from lemmata C.3 and C.2, any maximizer of  $\Phi^{\lambda}(\alpha, R)$  for any  $\lambda \in (0, 1)$  is non-increasing in  $\lambda$ . ■

**Proof.** of proposition 3.3. Recall that by lemma C.1, given a distribution  $F$  over projects the first order conditions for the problem are

$$\Phi_{\alpha}^F(\alpha, m) = -m \binom{R}{m} P^F(\alpha, m) V^F(\alpha, m), \quad (\text{D.3})$$

where  $P^F(\alpha, m)$  is strictly positive and  $V^F(\alpha, m)$  changes sign only once and the superscript refers to the corresponding probability distribution.

Assume that  $\alpha^F$ , the optimal cost sharing under  $F$  is in the interior,  $\alpha^F \in (1, R/m)$ .

Then  $V^F(\alpha^F, m) = 0$ .

Recall that

$$V^F(\alpha, m) = E_F(w(b) + \alpha | \alpha < b_1 < \bar{B}, \dots, \alpha < b_{m-1} < \bar{B}, \underline{B} < b_{m+1} < \alpha, \dots, \underline{B} < b_R < \alpha),$$

the latter can also be thought of as a function of the boundaries,  $h(\bar{B}, \underline{B})$ .

Recall that  $G$  is just a “shift” of distribution  $F$ , thus the corresponding value,  $V^G(\alpha, m)$ , under  $G$  is equal to

$$E_F(w(b) + \alpha | \alpha < b_1 < \bar{B} + a, \dots, \alpha < b_{m-1} < \bar{B} + a, \underline{B} + a < b_{m+1} < \alpha, \dots, \underline{B} + a < b_R < \alpha),$$

it can be represented as same function  $h$ , but evaluated at  $\overline{B} + a, \underline{B} + a$ . But, by theorem 5 from Milgrom and Weber (1982)

$$h(\overline{B} + a, \underline{B} + a) - h(\overline{B}, \underline{B}) \geq 0$$

It implies that

$$0 = V^F(\alpha^F, m) \leq V^G(\alpha^F, m),$$

and hence  $\alpha^F \geq \alpha^G$ . Corner cases are trivial. ■

## E. Appendix: A Counter-Example

### E.1. First Order Stochastic Dominance as a Criterion

In this example we'll illustrate that if one project first order stochastically dominates the other, it may be associated with *more differentiated* optimal tax, so that FOSD is not a suitable criterion (too weak) for our purposes.

It will be simpler to fix the ideas with a discrete distribution, although it can be translated into a continuous framework if necessary.

**Example E.1.** Consider the following case:  $R = 3, m = 2, \alpha = 1$ . Assume that conditional on  $b_2 = 1$ , joint distribution of  $(b_1, b_3)$  is  $h$ .

Distribution  $h$  :

benefit	$b_1$	$a - 3$	$a - 2$	$a + 2$	$a + 3$
$b_3$					
$a - 3$		0	$(1 - p)\varepsilon$	$\frac{1}{4}\delta$	$\frac{1}{2}\delta$
$a - 2$		$\frac{(1-d)}{2}\delta$	$p\varepsilon$	$\delta$	$\frac{1}{4}\delta$
$a + 2$		$\varepsilon$	$\delta$	$p\varepsilon$	$(1 - p)\varepsilon$
$a + 3$		$d\delta$	$\varepsilon$	$\frac{(1-d)}{2}\delta$	0

Let  $d = .99, p = .9, \varepsilon = 0.05, \delta = 0.2; a = (R - 1) / 2$ .

The table defines a probability density, indeed,

$$\delta(d + 2 + 1/2) + 2 \left( \varepsilon + p\varepsilon + (1/4)\delta + \frac{(1-d)}{2}\delta + (1-p)\varepsilon \right) = 1$$

The benefits are affiliated. Indeed,

$$\begin{aligned}
d\delta^2 - \varepsilon^2 &= 0.0371 > 0 \\
\delta^2 - p^2\varepsilon^2 &= 3.7975 \times 10^{-2} > 0 \\
p\varepsilon^2 - \frac{(1-d)}{2}\delta^2 &= 0.00205 > 0 \\
p\varepsilon\frac{1}{4}\delta - \delta(1-p)\varepsilon &= 0.00125 > 0 \\
d\delta^2 - \left(\frac{(1-d)}{2}\delta\right)^2 &= 3.9599 \times 10^{-2} > 0 \\
\frac{1}{2}\delta^2 - ((1-p)\varepsilon)^2 &= 1.9975 \times 10^{-2} > 0
\end{aligned}$$

In addition consider distribution  $g$ , which, conditional on  $b_2 = 1$  is described by the table below, and coincides with  $h$  otherwise

benefit	$b_1$	$a - 3$	$a - 2$	$a + 2$	$a + 3$
$b_3$					
$a - 3$		0	$(1-p)\varepsilon$	$\frac{1}{4}\delta$	$\frac{1}{2}\delta$
$a - 2$		$\frac{(1-d)}{2}\delta + (1-p)\varepsilon$	$p\varepsilon$	$\delta$	$\frac{1}{4}\delta$
$a + 2$		$p\varepsilon$	$\delta$	$p\varepsilon$	$(1-p)\varepsilon$
$a + 3$		$d\delta$	$p\varepsilon$	$\frac{(1-d)}{2}\delta + (1-p)\varepsilon$	0

In this case the benefits are affiliated as well,

$$\begin{aligned}
d\delta^2 - p^2\varepsilon^2 &= 3.7575 \times 10^{-2} > 0 \\
\delta^2 - p^2\varepsilon^2 &= \frac{1}{25} - \frac{1}{400}p^2 = 3.7975 \times 10^{-2} > 0 \\
p\varepsilon^2 - \delta\left(\frac{(1-d)}{2}\delta + (1-p)\varepsilon\right) &= 0.00105 > 0 \\
p\varepsilon\frac{1}{4}\delta - \delta(1-p)\varepsilon &= 0.00125 > 0 \\
d\delta^2 - \left(\frac{(1-d)}{2}\delta + (1-p)\varepsilon\right)^2 &= 3.9564 \times 10^{-2} > 0 \\
\frac{1}{2}\delta^2 - ((1-p)\varepsilon)^2 &= 1.9975 \times 10^{-2} > 0
\end{aligned}$$

It is clear that  $G$  first order stochastically dominates  $H$ .

Recall that the derivative of the objective function can be represented as

$$\Phi_\alpha(\alpha, m) = -m \begin{pmatrix} R \\ m \end{pmatrix} P(\alpha, m) V(\alpha, m), \quad (\text{E.1})$$

In this case under distribution  $H$

$$\begin{aligned}
V_H(1, 2) &= E_H(b_1 + b_3 + 1 - R | b_1 > 1, b_3 < 1) = \\
&= ((1-d)\delta(2-3) + p\varepsilon(2-2) + (1-p)\varepsilon(3-2)) = 0.003
\end{aligned}$$

Thus, objective function is decreasing for  $\alpha \geq 1$ , it follows that  $\alpha = 1$  is the solution to this problem (the constraint is binding).

Under  $G$ , however,

$$\begin{aligned} V_G(1, 2) &= E_G(b_1 + b_3 + 1 - R | b_1 > 1, b_3 < 1) = \\ &= (((1 - d)\delta + (1 - p)\varepsilon)(2 - 3) + p\varepsilon(2 - 2) + (1 - p)\varepsilon(3 - 2)) = -0.002, \end{aligned}$$

*It follows that the objective function is increasing at the boundary, thus the optimum in this case is strictly above unity.*

*Thus, it is optimal to set more “restrictive” tax (higher  $\alpha$ ) under  $G$  than under  $H$ , although  $G$  FOSD  $H$ .*