

DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 03-09

Can Decentralization Be Beneficial?

Anna Rubinchik-Pessach[†]

Revised: October 2003

Abstract

A conventional justification for government hierarchy in the fiscal federalism literature is based upon asymmetry in policy tools or in information access that is available to different levels of government. This paper demonstrates that even if these asymmetries are eliminated, addition of local (regional) governments to a one-tier central government can be strictly welfare improving.

Keywords: public goods, decentralization

JEL Classification Codes: H41, H72, H73, H11, D72

*This paper is based on my dissertation. I am indebted to Andrew Postlewaite for his guidance, and to Nicola Persico and Steve Matthews for insightful discussions. The paper was revised while the author was visiting CORE, the hospitality of which is greatly appreciated. The research was partially supported by the European Commission Network - HP RN-CT-2000-00072. I would also like to thank Ann M. Carlos, David Cass, Masa Fujita,

1 Introduction

Hierarchical governments, either federations or confederations, are widely spread across the world and are persistent across time. This may suggest that they are outperforming the unitary ones. However, according to a textbook economic argument, a central government can internalize the externalities, hence it has an advantage over a set of local ones, where politicians, representing their constituents, are motivated solely by local interests. So in a world, in which both levels of government have access to the same information, the role for local governments is unclear. Nevertheless, even in such a world maintaining a multi-leveled government may be worthwhile as opposed to a potentially cheaper unitary one. This paper provides an explanation.

Let me briefly describe the set up. Regional representatives, or legislators, striped off their "personal position" (say, identified by the district they will represent),

according to which the central legislature operates and its performance. If the designer is unconstrained, he can condition the constitution on the benefits and costs that become known to the legislators once a project appears on the agenda, and attain the ex-post optimum as in Laffont and Maskin (1982). Then there would be no need for local governments. This result, however, rests on the assumption that the realization of the payoffs is truly exogenous, in other words, it is not subject to strategic manipulations. Unfortunately, in many practical cases the assumption is violated, as estimating benefits and costs of a public project requires specific knowledge, or expertise, that politicians lack. For example, in the U. S. the evaluation of water projects is performed by the U. S. Army Corps of Engineers. Clearly, once it is done, predicted benefits and costs become common knowledge among the legislators.² In this case, conditioning the cost sharing arrangements on benefits would be equivalent to delegating the formulation of the rules to the "experts," who conduct the evaluation study. As their objectives may not coincide with those of the elected representatives, and, especially, those of the general electorate,³ this can result in additional informational rents extracted by the experts, which, clearly down-plays the attractiveness of such a

Gilboa, Postlewaite, and Schmeidler (2002)), the enforcement of such a constitution can be problematic, both because some contingencies are impossible to verify and also due to the intricacies of creating a proper motivation for the judges (see Tirole (1999) for the overview). Abstracting from the problem of formulating an “optimally incomplete” contract, which takes into account all the above mentioned considerations, I will follow Aghion and Bolton (2003) in assuming a particular form of incompleteness, so prevalent in practice. A public good has to be provided if at least m out of R legislators vote in its favor. In contrast to Aghion and Bolton (2003), taxes in this model can be specified ex-ante and can be conditioned on the voting behavior.

The core primitive of the model is the distribution, F , over future public projects as perceived by the constitutional designer. It reflects the likelihood of different profiles of gross benefits that will accrue from the projects. Regional variation in the willingness to pay for a project can result from a disparity in tastes, or from technological constraints. For example, the appeal of a publicly broadcasted ballet can vary across regions according to tastes, whereas the benefi

1.1 Related Literature

It would be almost impossible to trace the first discussion of the subject. Already by mid-19th century Guizot (1861),⁴ emphasized the tensions between a centrally exercised power, that is "...generally more disinterested, and more capable of taking justice and reason for its sole guide," and local institutions, the preponderance of which was attributed to the "infancy of societies," but which, are, nevertheless needed as the guards against usurpation of power by the center. Under a "truly representative" government described in this model every citizen enjoys sufficient protection from being discriminated against, at least on average, as the expected value of projects accepted by the central government is positive.

Later contributions focused on optimal allocation of responsibilities across levels of government based on the scope of public projects (Oates (1972), Lockwood (2002), Besley and Coate (2000) among others), access to information with respect to benefits and costs of the projects by level of government (Gilbert and Picard (1996), Zantman (2002), see Cr mer, Estache, and Seabright (1996) for an overview), or the access to the tax instruments available to the governments (Boadway, Marchand, and Vigneault (1998), Hochman, Pines, and Thisse (1995); see Wellisch (2000) for an overview).

The first two branches of the literature are most closely related to this model, and will be discussed in some more detail. Oates (1972) suggested that central governments make better decisions with respect to "global" public goods, which concern everybody in a country, whereas lower tier is better at providing the "local" public goods, the effect of which is mainly regional. By that argument, which rests on the presumption that central government has to provide the same amount of public good to every region and that the taxes have to be uniform, a country with heterogeneous population needs both levels of government. This is consistent with empirical findings by Panizza (1999), who shows that the degree of centralization is negatively correlated with the differentiation of tastes, where the ethnic fractionalization is used as a proxy for heterogeneity of tastes with respect to public goods.

More recent work by Lockwood (2002) and by Besley and Coate (2000)

trade-off between centralization and decentralization still remains. A common source of inefficiency on the central level in these models is a budgetary externality created by equal cost sharing. Lockwood (2002), demonstrates that under majority rule with uniform taxation the legislature rejects some desirable projects and accepts some undesirable ones. The author concludes that in the absence of externalities generated by public goods, decentralization is preferable for the country, in which residents are identical within a region. Besley and Coate (2000) reach a similar conclusion by allowing for heterogeneity of tastes within a region. In the presence of equal cost sharing, voters in any region have a motivation to elect to the central legislature a 'biased' representative, who is ready to accept more public projects than a median voter.

Another argument justifying government hierarchy rests on the common wisdom that local governments have superior information about citizens' preferences or technologies to be used in provision of public goods (this asymmetry is assumed away in the current paper). Restriction on information acquisition by central governments is questionable both in reality and on theoretical grounds, see Crémer, Estache, and Seabright (1996), although a central government may have fewer incentives to acquire the relevant information than the local governments.

The approach used in this paper also borrows from the mechanism design literature. Most recently, Palfrey and Ledyard (2002) rationalized referenda for big populations with independent valuations by comparing them to the optimal procedure that solves the corresponding designer's problem.

In this paper optimal solution is identified for some environments, which allows to rank performance of a hierarchical government versus the unitary one. In other environments, dominance argument is used to make the case for or against decentralization. Note also that this framework is explicitly constructed for a small number of decision makers (legislators) and the dependence in the valuations plays an important role in the analysis.

The rest of the paper is organized as follows. The next section outlines

governments appears worthwhile.

2 The model

There are $R \geq 3$ regions, R is odd. The residents of each region have identical attitudes towards public goods, but the attitudes across regions differ. Each region has one representative in the federal (central) legislature. The legislature has to decide whether to provide a the public project that appears on the agenda.

A project is associated with a profile of benefits, $b = (b$

The description of a one-tier government is the same, apart from the fact that the last stage 2 is omitted.

Throughout the rest of the paper the addition of the local governments to

2.1 Voting in the Central Legislature

Assume that the constitutional designer is creating the rules of acceptance of projects subject to a set of constraints. So, consider a project b that appears on the agenda.

Assumption VO Project is accepted if and only if at least m legislators vote for it;

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

Assumption AN Taxes are anonymous, i.e., a tax can not depend on the name of the region;

Assumption BB The sum of the tax payments should equal to the cost of the project;

Assumption SPM The Supporters of the project Pay (weakly) More than those who oppose the project.

All the assumptions LOCAL, TAX, AN, BB, SPM will be adopted for the rest of the analysis, and will be discussed in the conclusions.

Specification of the cost sharing rule By anonymity, taxes may differ only on the basis of a legislator's voting decision. So if $k \geq m$ legislators vote in favor of the project, the designer has to set two levels of taxes: $t(Y)$ and $t(N)$. Let the tax of the supporter be

$$t(Y) = \frac{R}{k}, \quad (1)$$

if $k < m$ legislators vote for the project. When the taxes are uniform, $\frac{R}{k} = 1$, so that every region pays the average cost of the project, unity. By assumption SPM, $\frac{R}{k} > 1$. In other words, $\frac{R}{k}$ measures the amount by which a supporter pays more than the average cost. By BB, anyone, who opposes the project is required to pay

$$t(N) = R - \frac{R}{k} = R \left(1 - \frac{1}{k}\right)$$

first element, v_1 , of this vector in order to determine the set of outcomes of

sharing vector will be denoted simply by
constitution should specify the parameters (

R. Therefore, the optimal con-

This leads us to two important conclusions. First, the presence of local projects with few externalities is not sufficient to justify the existence of local governments. Indeed, would these projects be prevalent on the agenda, the rules (constitutions) of the form $(\cdot, 1)$ could have provided an adequate

Lemma 3.1 Let $b_{[1]}$ be distributed with some continuous (non-degenerate) probability density function S on $[0, B]$, $B > R$. Let

$$\frac{b_{[1]}}{b_{[1]}}, k \in \{2, 3, \dots, R\} \quad (9)$$

be constant for any realization b of benefits.¹³ Let $\lambda = \frac{1}{b}$. Then $(\lambda, 1)$ with $\lambda = \frac{1}{b}$ solves problem (6), so that $Z(\lambda, 1) = 0$.

Proof Start with problem (5), which in this case simplifies to

$$\max_{\lambda \geq 1} \int_0^{\infty} \lambda^{-b_1 - 1} dS b_1, \quad (10)$$

admitting (multiple) solutions of the form,¹⁴

$$(\lambda(m), m), \quad \lambda(m) = \frac{b_{[1]}}{b} = 1. \quad (11)$$

Note that the last constraint is satisfied for at least one of the solutions, as $\lambda = \frac{1}{b}$

by the central legislature (and it will not, if $((R + 1) / 2) \neq 1$ in this case). It does not follow, however, that decentralization is necessarily needed to restore efficiency. An optimal voting rule and cost sharing arrangements for the upper level legislature can achieve the goal without an addition of another layer of government.

The last assertion will also be true for another environment, without the “knife-edge” assumptions, as in the previous example. To stress the comparison with the accepted view, assume that legislature has to deal with (primarily) local projects, thus, the benefits are concentrated. In particular, assume that the second highest benefit is most likely to be below the average cost. In addition, assume that a public project often generates some spillovers. The following lemma demonstrates that no matter how small are the spillovers, decentralization is not beneficial, if the benefits are sufficiently concentrated.

The idea behind the proof is quite simple. First, observe that the problem for the two-tier government (6) can be reduced to the first problem by restricting the benefits to the subset of the support, $[0, R] \cup [0, B]$. Indeed, any project with at least one benefit realization above R will be accepted in the presence of local government for any pair (τ, m) . So, the “constitution” aff

Proof First, note that for any $m' > 1$ expected value of the projects, W , under any rule (\cdot, m') is bounded by $R(R - 1)$. Next, let $E(w(b) | M(\cdot, 1)) = K$. As $\epsilon = R - 1$, K can be made strictly positive by choosing ϵ small enough, reducing the chance that $b_{[2]}$ will fall below ϵ . Pick ϵ such that

$$\epsilon < \frac{pK}{R(R - 1)}, \quad (12)$$

and $K > 0$. Therefore $W(\cdot, 1) = pK > W$

other project is local $i \in \mathcal{C}$ with benefits accruing to just one region.¹⁷ More precisely,

Proposition 3.3 Assume the vectors of benefits, b^H and b^L generate constant benefit ratios

$$b^i_r / b^i_1, \quad i = H, L; r = 1, \dots, R \quad (13)$$

and let b^H_1, b^L_1 be identically and independently distributed distributed S on $[0, B]$, with bounded density s . Assume $\tau = (1, 0, \dots, 0)$ and $\tau_2 > \bar{\tau}$. Let the projects L and H , appear with frequencies $p, 1 - p$ accordingly. Then for $p \in (0, 1)$, i.e., when both local (type L) and global (type H) projects appear on the agenda with positive probability the two-tier government is strictly welfare improving over the one-tier government.¹⁸

Proof First, let us show that it is possible to attain ex-post efficiency with the two tier-government. Next, we have to verify that this is impossible with a one-tier government, which will conclude the proof.

Indeed, denote by $V^*(p)$ the highest attainable welfare. If it were possible to condition the constitution on realization of benefits, the designer would have required $(\tau, m) = (R, 1)$ for local projects (tha9.31Tf0.00y

same for both types of projects: H and L. Therefore, the highest attainable welfare in this case is

$$\max_{\epsilon \in [1, \infty)} (1 - p) \int_{\epsilon}^{\infty} \frac{z}{b_1 - 1} dS(b_1) + p \int_{\epsilon}^{\infty} \frac{z}{R} b_1$$

be at least two without the fear of losing good "local" projects, that can be picked up by the lower tier. This option is, clearly, infeasible for the unitary governments.

The proof of the next proposition is conceptually dissimilar to the previous one though. Without being able to ensure (as in the previous example) that the "constitution" maximizes ex-post welfare for $m > 1$, and, thus, can not be improved upon, one has to verify that the choice $m = 1$ is dominated. This is done by making the ability to condition on $b_{[2]}$ sufficiently attractive. I have introduced a parameter, θ , which could be interpreted as "news," or "general state of affairs" that the rules may depend upon. As θ decreases (meaning "bad news"), the ex-ante value of a project reduces (becomes more

that regional benefits are affiliated is consistent with the idea that a public project has an 'objective' (common) value perceived differently by regions. 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.

Proposition 3.5 Assume

1. F is symmetric in all its arguments with the corresponding marginal distribution f , f is continuous on $[0, B]$.
2. Assume random variables $b_k, k = 1, \dots, R$ are affiliated;
3. Let $v : R \times R \rightarrow R; v(x, x) = E \sum_i w(b_i) | b_{[1]} = x; v(x, x) < v(x, y) < v(y, x) < v(y, y)$ is strictly increasing in $x \in [0, 1]$ for all $x \in [0, R]$;
4. Let $z(x) = E \sum_i w(b_i) | b_{[1]} = x, t(x) = E \sum_i w(b_i) | b_{[2]} = x, z(x) < 0 < t(x) < v(x, x)$ for all x .

Then there exist a threshold $\bar{x} < \underline{x}$, such that for any $x < \underline{x}$, welfare attained under decentralization (in problem (6)) is strictly higher than that under unitary government (in problem (5)).

Proof is relegated to the appendix. ■

4 Conclusions

This model provides an economic rationale for the existence of a hierarchical government, which may outperform the unitary one even in the absence of the vertical asymmetric information within the hierarchy and even though the absolute advantage of the (upper) central government in making decisions with respect to public goods is evident. Somewhat contrary to the accepted view, if the issues to be tackled by the government are predominantly of local importance, albeit generating some spillovers, central government alone is sufficient, i.e., there is no need to decentralize. On the other hand, it is the prevalence of valuable global issues along with the projects of local importance that generates the need for the local governments. The hierarchy induces specialization of each tier on the corresponding issues, thus, enhancing the overall welfare, which can justify possible costs associated with the

additional level of government. Thus, the main argument does not stem from an assumed deficiency of a central government, but rather, rests on the idea of specialization.

Let me briefly discuss the assumptions that underlie the conclusion. First, the restriction to the yes-no voting is crucial. Clearly, departing from voting (extending the set of the possible messages) can restore efficiency, see Laffont

erates under the majority rule with uniform taxation ($m = 2$, $n = 1$). Clearly, the second project, L , will be accepted by the first region (on the local level), as $b_1 = 4 > 3$.

- Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4, 25–55.
- Tirole, J. (1999). Incomplete Contracts: Where Do We Stand? *Econometrica* 67 (4), 741–781.
- van Dorp, J. and S. Kotz (2002). The Standard Two Sided Power Distribution and its Properties: with Applications in Financial Engineering. *The American Statistician* 56, 90–99.
- Wellisch, D. (2000). *Theory of Public Finance in a Federal State*. Cambridge University Press.
- Zantman, W. (2002). Constitutional Design and Regional Favoritism. *Journal of Public Economic Theory* 4, 71–93.

A Appendix

Proof of proposition 2.

standard trembling hand perfection argument rules out equilibria supporting
 $m - k$

Proof By symmetry of F, objective function $W(\mu, m)$ for a given $R, m \geq 1$, can be represented as $\Phi(\mu, m)$,

$$\Phi(\mu, m) = \prod_{k=1}^R \mu_k \prod_{i=1}^m |z_i - z_{i-1}| \dots |z_1 - z_0|$$

3

Proof Let

$$V(x, m) = \frac{1}{P(x, m)} \int_{|z|^{-1}}^x \int_{|z|^{-1}}^z \int_{|z|^{-1}}^z \int_{|z|^{-1}}^z [w(b_{-1}, z)] dF(x, b_{-1}), \quad (23)$$

$$w(b_{-1}, z) = x - b_{-1} - R, \quad (24)$$

$$P(x, m) = \int_{|z|^{-1}}^x \int_{|z|^{-1}}^z \int_{|z|^{-1}}^z \int_{|z|^{-1}}^z dF(x, b_{-1}), \quad (25)$$

$$b_{-1} = (b_1, b_2, \dots, b_{-1}, b_{+1}, \dots, b). \quad (26)$$

Clearly, then

$$P(x, m) V(x, m) = A(x, m),$$

where $A(x, m)$ is as defined in 18, which justifies 22. It is obvious that $P(x, m) > 0$ for $x \in [1, R)$.

Therefore, it is left to show that $V(x, m)$ is strictly increasing in the first argument. Indeed, in the view of definition 23,

$$V(x, m) = E \int_{|z|^{-1}}^x w(b_{-1}, z) |b_{-1}| \dots |b_{+1}| dF(x, b_{-1}) = \int_{|z|^{-1}}^x \int_{|z|^{-1}}^z \int_{|z|^{-1}}^z \int_{|z|^{-1}}^z [w(b_{-1}, z)] dF(x, b_{-1})$$

for any feasible λ' . Provided $Z(\lambda, 2) > 0$ for any λ , it is sufficient to show that

$$W(\lambda, 2) - W(\lambda', 1) \geq 0. \quad (29)$$

conditional on $b_{11} \in \mathbb{R}^{20}$ in other words, $(\lambda', 1)$ can not be a solution to problem (5) for any λ' .

is increasing in the first two arguments by Milgrom and Weber (1982), theorem 5. Let us focus on the non-degenerate case of t being strictly increasing in x and y .²² Note that v is also continuous in x and y due to continuity of f .

Let $v^*(x) = v(x, y^*)$. Need to show that $v^*(x) - t(x) > 0$ for low enough x . Clearly, $v(x, R) > t(x)$ by definition of $t(x)$. Next, it is easy to verify that $v(x, y^*) < 0$, therefore, $v(x, y^*) < t(x)$, as $t(x) > 0$ for all x by assumption. Fix $x = \hat{x}$. Then there exists $y < R$, such that

$$v(\hat{x}, y) = t(\hat{x}). \quad (37)$$

If $y < y^*$, then, clearly, $v(\hat{x}, y) = v^*(\hat{x}) - t(\hat{x}) > 0$. If $y > y^*$, fix $\hat{x} = x$, and $\hat{t} = t(\hat{x})$ then decrease x to x_0 . As a result y should increase. Indeed, y^* solves

$$V(x, 1; y) = 0; \quad (38)$$

$$V(x, 1; y) = v(x, y) = E^i w(b_{-1}, y) | = b_{[1]}; \quad (39)$$

and $v(x, y)$ is assumed to be strictly increasing in y . It follows that y^* is decreasing in x . Therefore, for fixed $t = \hat{t}$ and $x = \hat{x}$ inequality $y < y^*$ is easier to satisfy as x decreases. In addition, given t is strictly increasing in the first two arguments, and in the view of equation 37, y should increase with t and decrease with

the following equation,²⁴

$$E^i w(b) | b_{2j}$$