A Principal-Agent Building Block for the Study of Decentralization and Integration

Mariano Tommasi
CEDI (Fundación Gobierno & Sociedad)
Universidad de San Andrés
tommasi@udesa.edu.ar

Federico Weinschelbaum[1]
Universidad de San Andrés
fweinsch@udesa.edu.ar

This Version: November 1999
Comments Welcome

Abstract

The architecture of public decision making in the world is being changed through processes of "economic integration" and of "decentralization". Some policy decisions are now taken at a higher level (i.e., monetary policy in Europe, trade policy in part of South America), while others are taken by smaller political units "closer to the people" (i.e., health and education policies in many Latin American countries). We provide a building block for the study of such processes, emphasizing the trade-off between the advantages of centralized decision making (internalization of externalities) and those of decentralized decision making (increased principal-agent control by the citizens). We do so within the context of a class of principal-agent models known as common agency.

JEL CLASSIFICATION NUMBERS: D62, D71, H49, H79.
KEYWORDS: centralization, integration, agency, externalities.

[1]We thank the research assistance of Gabriel Basaluzzo, and the valuable comments of Alex Galetovic, Matías Iarykzower, Pablo Spiller, Ernesto Stein, Marcelo Veracierto and seminar participants at Universidad de San Andrés, at the Meeting of the Latin America and the Caribbean Economic Association in Santiago de Chile and at the Annual Conference of the Central Bank of Uruguay.

1 Introduction

The architecture of public decision making in the world is being dramatically altered through processes of "integration" and of "decentralization." Some policy decisions are now taken at a higher level (i.e., monetary policy in Europe, trade policy in part of South America), while others are taken by smaller political units "closer to the people" (i.e., health and education policies in many Latin American countries).

In a sense, both processes are the two faces of the same coin. When it is deemed beneficial to provide public goods (or to make policy decisions) at a higher level of aggregation, we speak about integration. When it is considered that the provision of some public goods is better located at a lower level than the current one, we speak about decentralization.

The objective of this paper is to contribute to a growing body of literature (loosely, on "federalism") analyzing these issues, and to attempt to identify some of the characteristics of the goods, policies or circumstances that call for provision at different levels in a governmental hierarchy.[2]

In particular, we want to formalize the trade offs between some advantages of centralized decision making - namely, the internalization of externalities and economies of scale - and some of its disadvantages - namely the "democratic deficit" of having decision making further removed from the citizenry. To focus on this latter point, we believe that it is necessary to remove the assumption of benevolent governments, and to utilize a principal agent framework.[3] The rhetoric of current decentralization efforts (see, for instance, World Bank 1999) emphasizes notions such as accountability, proximity, yardstick competition, all of which should, in our view, be castled in formal agency set-ups. It might be the case that, with larger and more dispersed populations, it is harder to solve the free-rider and coordi-

[2] Besley and Coate (1998) provide an excellent overview of that literature, as well as making a particular contribution.

[3] We share the spirit of Qian and Weingast (1997) who also call for opening the black box of political organizations in the study of federalism.

nation problems that arise in controlling "the agent" we call the government. In that sense, decentralization (bringing government closer to the people) might be a way of alleviating political control problems.

As a first step in this agenda, for the sake of generality and comparability with other areas of application, we cast our analysis in a relatively standard principal-agent framework. We analyze a case in which the principal is not a single individual but a group, a population. The core of the problem is the interaction inside this "collective principal". To our surprise, the "collective principal" case is a piece of the agency literature which is not fully developed. We have borrowed the "common agency" model (Bernheim and Whinston 1986, Grossman and Helpman 1994, Dixit, 1996) to give a first cut to the study of the problem. One variant of the common agency model, known as intrinsic common agency is a good first approximation to the problem of control of policymakers by citizens. We discuss the general class of multiprincipal-agent models and its applicability to our problem in later sections.

Our model has two essential ingredients: an externality problem in the provision of ("local") public goods (favoring centralization as the desired institutional arrangement), and a collective action problem among (citizens) principals in controlling political agents (favoring decentralization). The first component has been a standard feature in the discussion of the trade-offs between centralized and decentralized provision of public goods since, at least, the seminal work of Oates (1972). In that paper, the externality/spillover effect was traded-off against the cost of centralized provision in terms of a \one size fits all" policy of uniform public good provision, independently of local needs and tastes. Oates' Decentralization Theorem states that in the absence of spillovers (and of cost-savings from centralized provision), decentralization is preferable. This has to be read as \preferable to uniform provision." But, in a setting of perfect information, nothing will prevent a benevolent central planner to prescribe the right amounts for each jurisdiction.

Later work has emphasized, hence, that the case for decentralization has to be driven by political economy considerations. Besley and Coate (1998), Lockwood (1998) and Seabright

(1996) present models in which potential benefits of decentralization are derived through endogenous choices under alternative political aggregation mechanisms. Bardhan and Mookherjee (1998) analyze alternative methods of delegating authority, a central government has limited ability to monitor the performance of the bureaucrats while in a decentralized system the local governments may be subject to capture by local elites.

Many of those papers, do require interjurisdictional heterogeneity "a la Oates" in order to derive benefits of decentralization. One of the features of our formalization is that it does not require heterogeneity. In the simplest formulation of the heterogeneity issue, decentralization can improve the efficiency of governments because local officials have better information to match the mix of services produced by the public sector and the preferences of the local population (i.e., they have the means to be responsive). The principal-agent avenue that we pursue emphasizes the incentives of politicians to better serve their people. We believe that our model provides a useful step in the process of formalizing some of the key concepts being discussed in the decentralization debate around the globe.

2 The Model[4]

There are M towns. A local public good has to be provided for each town. Hence, we have an M goods economy $x = (x_1; x_2; :::; x_M)$: There are $N = n_1 + n_2 + :::+ n_M$ citizens (principals) of type $1; 2; ::M$ respectively.

We assume that each principal has linear preferences according to his type

$$u_i(x) = b_{i1} \cdot x_1 + b_{i2} \cdot x_2 + :::+ b_{iM} \cdot x_M = b_i^0 \cdot x$$

$b_{ii} > 0$ is the utility that each principal of type i gets for a unit of his own local public good and $b_{ij} > 0$ $(i \neq j)$ is the externality that he gets for a unit of local public good in town j.

We will consider two alternative "federal" organizations; one in which there is one agent serving the whole population, and another in which there is one agent per locality. (In the

[4] We follow the formulation of the common agency model of Dixit (1996).

second case, "decentralization," we do not allow contracting between citizens in one locality and policymakers in another.

The production technology is given by a level of "effort" (t_i) chosen by the agent responsible to provide the local public good for each town plus an error term (ε_i). The error terms are independently and normally distributed with mean 0 and variance σ_i^2. (In the more general case there will be a variance matrix - which might include non-zero off-diagonal elements.)

$$x = t + \varepsilon;$$

where t is the vector of the agent(s)' efforts, $t = (t_1; t_2; ...; t_n)$, and $\varepsilon \in R^M$ is the vector of error terms.

As common in the principal-agent literature, agents are risk averse. We assume that they have constant absolute risk aversion, with utility function

$$u_a(w) = -e^{-rw};$$

where w is the monetary measure of the utility and is composed by the payment z that they receive from the principals minus a quadratic cost of effort $\frac{1}{2}t'Ct$ where[5]

$$C = \begin{bmatrix} c_1 & 0 & 0 & \cdots & 0 \\ 0 & c_2 & 0 & \cdots & 0 \\ 0 & 0 & c_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & c_m \end{bmatrix}:$$

Hence when there is only one agent, his payoff is

$$w = z - \frac{1}{2}t'Ct = z - \frac{1}{2}\sum_{j=1}^{n} c_j t_j^2;$$

[5] The assumption of C being a diagonal matrix rules out the possibility of having externalities in the production side.

and when there are M agents, their payoffs are

$$w_i = z_i - \frac{1}{2} t_i^2 c,$$

with $z = \sum_{i=1}^{M} z_i$.

Principal i expected utility is $\sum_{j=1}^{N} b_{ij} t_j - z_i$. The "aggregate" principal receives, in expected value[6]

$$\sum_{j=1}^{N} \left(\sum_{i=1}^{M} n_i b_{ij} \right) t_j - z.$$

In the remainder of this section, we evaluate the welfare that is attained under two alternative institutional arrangements: centralization, when the whole population hires one agent to provide the whole vector of goods, and decentralization, when each town hires its own agent to provide the local public good. We do so under three different contexts in terms of observability of the agents effort and in terms of the nature of interactions among principals. In subsection 2.1., effort is observable and verifiable (hence contractable) and the principals act as unified actors – there is no problem of cooperation among principals in contracting with the agent. In subsection 2.2., we maintain the assumption of united principals, but effort is not observable. Finally, in subsection 2.3., effort is not observable and principals act in a non-cooperative manner. The first two cases serve as benchmark for the third one, the one we focus on.[7]

As it is a standard practice in these models, we assume that the principal offers a contract and the agent can accept or reject it, implicitly giving all the bargaining power to principals. (There are some subtleties in applying this logic to common agency cases. We refer to that in 2.3.) When effort is not observable, the agent, after signing the contract, will decide the level of effort that he will offer.

[6] We are using the notation j to refer to goods, and i to reserve to principals' type.

[7] There is a fourth context we do not analyze explicitly: that of contractable effort and separate principals. It is easy to see that in that case, results analogous to contractable effort with united principals exist (see Bernheim and Whinston, 1986).

2.1 Observable and Verifiable Effort, United Principals

In this case principals and agents can write contracts contingent on the agents providing a stipulated level of effort.

2.1.1 Centralized case

Since the payment is only a transfer and it will be at the level that gives to the agent his reservation utility, the principal(s) will choose the level of effort that maximizes aggregate surplus,

$$\sum_{j=1}^{N}\left[\sum_{i=1}^{M} n_i b_{ij} t_j - \frac{1}{2} c_j t_j^2\right] \tag{1}$$

The first order condition with respect to t_j, leads to

$$\sum_{i=1}^{M} n_i b_{ij} = c_j t_j .$$

Marginal social benefit is equated to marginal social cost. For this centralized case, as usual in principal-agent models, first-best is achieved when effort is contractable. The level of effort is

$$t_j^{COU} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_j} = t_j^{*}$$

for all j, where COU stands for (Centralized, Observable, United), and $*$ stands for the socially optimal level.

We will use this case not only to compare it with the decentralized one, but also as a benchmark to compare with other environments. Since the aggregate surplus is a quadratic function on t_j that achieves a maximum when $t_j = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_j}$, we know that if $t_j < \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_j}$; t_j is a measure of welfare.[8]

[8] This is clearly valid when the agents' payment is riskless, as in this case. When effort is not observable, contracts will be such that agents' will bear some risk, and social surplus will have a term in addition to those in equation (1) to capture that loss. We will show that the claim of t being a sufficient statistic for welfare will still be valid in that case.

2.1.2 Decentralized case

Now we have separate agents. Their respective costs are $\frac{1}{2}c_i t_i^2$.

Since the payment that the aggregate principal of each region gives to his agent is only a transfer and it will be at the level that gives to the agent the reservation utility, this aggregate principal will choose the level of effort that maximizes the aggregate surplus of the locality.

Type i principals maximize $\sum_{j=1}^{M} n_i b_{ij} t_j - \frac{1}{2}c_i t_i^2$ with respect to t_i; taking t_k $(k \neq i)$ as given, leading to $n_i b_{ii} = c_i t_i$, that is the marginal social cost equals the marginal social benefit of the locality. Although effort is contractable the result is not optimal since each principal does not take into account the externalities that its good provides to the other regions. The level of effort in a decentralized world is

$$t_i^{DOU} = \frac{n_i b_{ii}}{c_i} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_i} \frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}} = t_i^{*} \frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}}.$$

We will have under-provision of effort.

Therefore, centralization is better due to the spillovers, which in this set up we might want to call consumption externalities. With positive production externalities ($c_{ij} < 0$) the effect would be magnified. If we introduce now a parameter δ reflecting the average size of spillovers, such that $u_i(x) = b_{ii} x_i + \delta \sum_{j \neq i} b_{ij} x_j$, we see that the departure from the centralized solution is $\Delta t_i = \frac{\delta \sum_{j \neq i} n_j b_{ij}}{c_i}$, which reduces to zero for the case of pure private (or pure local public) goods ($\delta = 0$), and reaches a maximum for the case of pure global public goods ($\delta = 1$).

2.2 Non Observable Effort, United Principals

2.2.1 Centralized case

Now principals can monitor the efforts t but only imperfectly, i.e., they observe x but not t. The agent is hired through contracts with payments

$$z = \alpha' x + \beta.$$

We follow Dixit (1996) in restricting attention to linear reward schemes, since they go naturally with quadratic payoffs (see also Holmstrom and Milgrom 1987 and 1991). α is the vector of incentive payments for each activity, while β is used to fulfill the agent's participation constraint.

The agent's certainty equivalent (CE) is $\alpha' t + \beta - \frac{1}{2} r \alpha' \Omega \alpha - \frac{1}{2} t' C t$; and the principals benefit is $\sum_{j=1}^{n} \left(\sum_{i=1}^{m} \pi_i b_{ij} - \alpha_j \right) x_j - \beta$.

The natural equilibrium concept is Perfect Bayesian Equilibrium, hence we start by solving for the effort choices of the agent, which will lead to his incentive compatibility constraint. The agent's problem is

$$\max_t \sum_{j=1}^{n} \alpha_j t_j - \frac{r}{2} \sum_{j=1}^{n} \alpha_j^2 \sigma_j^2 - \frac{1}{2} \sum_{j=1}^{n} c_j t_j^2 + \beta$$

which leads to

$$t_j = \frac{\alpha_j}{c_j}$$

for all j.

The expected social surplus is then

$$\sum_{j=1}^{n} \left[\sum_{i=1}^{m} \pi_i b_{ij} \left(\frac{\alpha_j}{c_j} \right) - \frac{r}{2} \alpha_j^2 \sigma_j^2 - \frac{1}{2} c_j \left(\frac{\alpha_j}{c_j} \right)^2 \right] \tag{2}$$

The principal maximizes (2) with respect to α leading to $\alpha_j^{CNU} = \frac{\sum_{i=1}^{m} \pi_i b_{ij}}{(1 + c_j r \sigma_j^2)}$, which implies

$$t_j^{CNU} = \frac{\sum_{i=1}^{m} \pi_i b_{ij}}{c_j} \left(\frac{1}{1 + r c_j \sigma_j^2} \right) = t_j \left(\frac{1}{1 + r c_j \sigma_j^2} \right)$$

The level of effort is lower than the optimal whenever the coefficient of absolute risk aversion (r) is positive. This is the "traditional" principal agent problem, with its associated trade off between incentives and risk sharing.

Note that the expected social surplus is quadratic in α_j; then α_j is a measure of welfare when $\alpha_j < \alpha_j^{CNU}$, and so is t_j since it is increasing in α_j; when $t_j < t_j^{CNU}$. Hence, as in the case when effort was contractable, t is a measure of welfare.

2.2.2 Decentralized case

The problem of the agents, and its solution, is the same as the one of the centralized case.

Type i principals maximize

$$\sum_{j=1}^{M} n_i b_{ij} \frac{\alpha_j}{c_j} - \frac{r}{2}\alpha_i^2 \sigma_i^2 - \frac{1}{2} c_i \left(\frac{\alpha_i}{c_i}\right)^2$$

with respect to α_i taking α_k $(k \neq i)$ as given, obtaining $\alpha_i^{DNU} = \frac{n_i b_{ii}}{(1 + c_i r \sigma_i^2)}$, which implies

$$t_i^{DNU} = \frac{\sum_{j=1}^{M} n_i b_{ij}}{c_i} - \sum_{j=1}^{M} n_i b_{ij} \frac{1}{(1 + c_i r \sigma_i^2)} = t_i^{*} - \sum_{j=1}^{M} n_i b_{ij} \frac{1}{(1 + c_i r \sigma_i^2)} \quad (3)$$

The effort exerted by the agent is, again, lower for this case than for the centralized economy, except when consumption externalities are zero. As it can be seen in equation (3), in this case there are two sources of the under provision of effort, the externalities ($n_i b_{ii} < \sum_{j=1} n_i b_{ij}$) that the principals do not take into account, and the low power incentive scheme that is given to the agents in order to minimize their exposure to risk.

Note that in the two cases considered so far, centralization is preferable. That is because we haven't brought into play yet the potential disadvantage of centralization, in terms of a harder agency problem due to the larger number of principals. To that we turn now.

2.3 Non observable Effort, Separate Principals

This is a good point to pause and to pinpoint which is the exact exercise to be performed, its antecedents in the theoretical literature, and its relevance for the applied problem at hand. We are studying situations in which there are multiple principals (citizens, in the specific application). The interactions among principals in their relation to the agent/s might operate at different stages of a full blown principal agent interaction. In this paper, we follow the precedent of Bernheim and Whinston (1986), Grossman and Helpman (1994), Dixit (1996), Dixit, Grossman and Helpman (1997) and Stole (1997)[9] in their focus on the

[9] Our model is closer to the first three (it is literally an extension of Dixit 1996) in focusing on the moral hazard case with noncontractible effort. Dixit et al (1997) treat the case in which effort is contractible (with general, as opposed to quasi-linear, preferences). Stole (1997) focuses on an adverse selection case.

contracting stage of the control problem. Other stages of control, such as monitoring or, perhaps, enforcement can be also subject to the collective action problems in control which seem to underlie the conventional wisdom of "small is beautiful" in the applied decentralization literature.

These papers have christened the problem they analyze as the "common agency" framework. In a nutshell, the games being analyzed consist of:

1) A first stage, in which each principal offers a payment scheme (a contract) to the agent

2) A second stage in which the agent decides whether to accept the contract/s

3) A final stage in which the agent (if, in stage 2, he decided to participate) decides upon his level/s of effort.

All of the papers maintain the assumption that the principal/s make take-it-or-leave-it offers to the agent. Interestingly, in this multiprincipal set up it is not so obvious that this implies that all the surplus of the relationships goes to the principal/s.[10] It turns out that the second stage, not always made explicit in the previous papers, might also impinge upon the "bargaining" outcome.

We can imagine two modelling choices at that point: either the agent is forced to choose between accepting or rejecting the full set of contracts offered, or he can choose whether to accept or reject each particular contract (which is equivalent to selecting any subset of contracts). Stole (1997) refers to the first case as intrinsic common agency and to the second as delegated common agency.[11] An example of intrinsic common agency would be

[10] As a matter of fact, Dixit et al (1997) show, for a particular case, that all the surplus will go to the agent as long as there is more than one principal.

[11] We will follow his definition since it is clear and it is exhaustive, even though we think it is not a completely satisfying terminology. Stole (1997) claims to be following the terminology of Bernheim and Whinston (1986), but we don´t think that is correct. In defining these two categories BW refer to whether the principal/s choose(s) to delegate some decisions to an agent, delegated agency, versus cases in which the agent is naturally endowed with the right to make some decisions which affect the welfare of the principal/s, intrinsic agency. Notice that, unlike Stole's their definition applies even to bilateral agency situations. Stole calls intrinsic common agency (a good name) cases in which the agent does not have the option to

regulation by multiple authorities: the regulated …rm can choose between abiding to all of those payment functions or exiting the market altogether.. A n example of "delegated" common agency would be a salesman who decides whose products to carry. A ll of the mentioned papers have modelled the intrinsic common agency case, and so do we because we believe it is the one that comes closer to capturing the problem of political control we want to study.

Coming to the modelling details, we consider our two standard cases. In the centralized case, each of the N total principals will o¤er a contract to the only agent which might be a function of the realized level of output in all the activities, as in the previous literature. In a N ash equilibrium, each principal o¤ers his contract, taking all the other principal's contracts as given.

In the decentralized case each principal in town i is allowed to o¤er a contract to the agent of that town but not to agents from other towns. W ithin each town, each principal o¤ers his contract, taking all the other principal's contracts as given. O ne way of thinking about our exercise would be to imagine that in the centralized case there are still M agents, but that "cross-contracts" are allowed and that those M agents are united.

2.3.1 Centralized case

In this case each of the $N = \sum_{i=1}^{M} n_i$ principals can contract with the one agent. P rincipals can monitor e¤orts t only imperfectly. T he agent is hired through individual contracts with each principal l and payment $z^l = \alpha^l x + \beta^l$; where $\alpha^l = (\alpha_1^l; \alpha_2^l; ::: ; \alpha_M^l)$: Each principal will o¤er a contract, taking as given the contracts o¤ered by the other principals. L et $\alpha_j = \sum_l \alpha_j^l$ and $\beta = \sum_l \beta^l$: T he agent's CE is $\alpha' t + \beta - \frac{r}{2}\alpha' \Omega \alpha - \frac{1}{2} t' C t$:

select a subset of principals and delegated common agency (a bad name) cases in which the agent can select principals.

At the stage of choosing effort, the agent maximizes

$$\max_{t} \sum_{j=1}^{N} \alpha_j t_j - \frac{r}{2} \sum_{j=1}^{N} \alpha_j^2 \sigma_j^2 - \frac{1}{2} \sum_{j=1}^{N} c_j t_j^2 + \varepsilon$$

which leads to $t_j = \frac{\alpha_j}{c_j}$.

Let $A_j^l = \sum_{k \neq l} \alpha_j^k = \alpha_j - \alpha_j^l$; and $B^l = \sum_{k \neq l} \varepsilon^k = \varepsilon - \varepsilon^l$. If only principal l does not sign a contract with the agent regarding activity j, the latter best strategy over j will be $t_j = \frac{A_j^l}{c_j}$, and his certainty equivalent (CE) will be

$$\sum_{j=1}^{N} \left[\frac{A_j^l}{c_j} - \frac{r}{2} \left(A_j^l \right)^2 \sigma_j^2 - \frac{1}{2} \frac{\left(A_j^l \right)^2}{c_j} \right] + B^l = \sum_{j=1}^{N} \left(A_j^l \right)^2 \left[\frac{1}{2 c_j} - \frac{r}{2} \sigma_j^2 \right] + B^l$$

whereas if this additional contract is signed, effort will be $t_j = \frac{\alpha_j}{c_j}$, and the agent's CE will be

$$\sum_{j=1}^{N} \left(A_j^l + \alpha_j^l \right)^2 \left[\frac{1}{2 c_j} - \frac{r}{2} \sigma_j^2 \right] + B^l + \varepsilon^l .$$

The marginal gains in signing the contract are

$$\sum_{j=1}^{N} \left[\left(\alpha_j^l \right)^2 + 2 A_j^l \alpha_j^l \right] \left[\frac{1}{2 c_j} - \frac{r}{2} \sigma_j^2 \right] + \varepsilon^l . \tag{4}$$

Principal l's expected utility if he does not sign the contract is $\sum_{j=1}^{N} b_{ij} \frac{A_j^l}{c_j}$ (where i is the town where citizen l lives), whereas if he signs the contract it will be $\sum_{j=1}^{N} \left[b_{ij} - \alpha_j^l \right] \frac{A_j^l + \alpha_j^l}{c_j} - \varepsilon^l$, and the marginal gains are

$$\sum_{j=1}^{N} \left[b_{ij} \frac{\alpha_j^l}{c_j} - \alpha_j^l \frac{A_j^l + \alpha_j^l}{c_j} \right] - \varepsilon^l . \tag{5}$$

From (4) and (5), the total bilateral surplus is then

$$\sum_{j=1}^{N} \left[b_{ij} \frac{\alpha_j^l}{c_j} - \alpha_j^l \frac{A_j^l + \alpha_j^l}{c_j} + \left[\left(\alpha_j^l \right)^2 + 2 A_j^l \alpha_j^l \right] \left[\frac{1}{2 c_j} - \frac{r}{2} \sigma_j^2 \right] \right] . \tag{6}$$

Maximizing (6) with respect to α_j^l leads to

$$0 = \frac{b_{ij}}{c_j} - \frac{A_j^l + 2 \alpha_j^l}{c_j} + \left[\alpha_j^l + A_j^l \right] \left[\frac{1}{c_j} - r \sigma_j^2 \right] =$$

$$= b_{ij} - \alpha_j^l - \alpha_j^l + A_j^l - c_j r \sigma_j^2 :$$

After solving for α_j^I we get $\alpha_j^I\left(1+r\sigma_j^2\right) = b_{ij} - A_j^I r\sigma_j^2$. Recalling that $A_j^I = \alpha_j - \alpha_j^I$ we obtain $\alpha_j^I = b_{ij} - \alpha_j r\sigma_j^2$. Adding over all the principals we obtain

$$\alpha_j = \sum_{i=1}^{N} \alpha_j^I = \sum_{i=1}^{N} n_i b_{ij} - N\alpha_j r\sigma_j^2 .$$

Therefore, for this case we have

$$\alpha_j^{CNS} = \frac{\sum_{i=1}^{N} n_i b_{ij}}{1+Nr\sigma_j^2} .$$

This gives a level of effort

$$t_j^{CNS} = \frac{\sum_{i=1}^{N} n_i b_{ij}}{c_j}\frac{1}{1+r\sigma_j^2}\frac{1+r\sigma_j^2}{1+Nr\sigma_j^2} = t_j^{*}\frac{1}{1+r\sigma_j^2}\frac{1+r\sigma_j^2}{1+Nr\sigma_j^2}; \tag{7}$$

which is smaller than in the case of united principals. The first term in the right hand side of (7) is the optimal value of t_j but this is multiplied by $\frac{1}{1+r\sigma_j^2}$ the "risk sharing" effect and by $\frac{1+r\sigma_j^2}{1+Nr\sigma_j^2}$ the "collective principal" effect.

2.3.2 Decentralized case

In this case we have that in each town i, the n_i principals will be offering contracts to the local agent, but not to agents in other localities. The marginal gain for agent i in signing a contract with principal I of his town is $\left(\alpha_i^{I\,2} + 2A_i^I\alpha_i^I\right)\left(\frac{1}{2c_i} - \frac{r}{2}\sigma_i^2\right) + \tau^I$.

Principal I's expected utility if he does not sign the contract is $\sum_{j=1}^{N} b_{ij}\frac{A_j^I}{c_j}$, whereas if he signs the contract it will be $\sum_{j=1}^{N} b_{ij}\frac{A_j^I+\alpha_j^I}{c_j} - \alpha_i^I\frac{A_i^I+\alpha_i^I}{c_i} - \tau^I$. The total bilateral surplus is then

$$\sum_{j=1}^{N} b_{ij}\frac{A_j^I+\alpha_j^I}{c_j} - \alpha_i^I\frac{A_i^I+\alpha_i^I}{c_i} + \left(\alpha_i^{I\,2} + 2A_i^I\alpha_i^I\right)\left(\frac{1}{2c_i} - \frac{r}{2}\sigma_i^2\right) : \tag{8}$$

Maximizing (8) with respect to α_i^I leads to

$$0 = b_{ii} - \alpha_i^I - \left(\alpha_i^I + A_i^I\right)r\sigma_i^2 :$$

After solving for α_i^I we obtain $\alpha_i^I(1+r\sigma_i^2) = b_{ii} - A_i^I r\sigma_i^2$. Recalling that $A_i^I = \alpha_i - \alpha_i^I$ this leads to $\alpha_i^I = b_{ii} - \alpha_i r\sigma_i^2$. Adding over all the principals that can contract agent i, i.e., over citizens of town i, we obtain $\alpha_i = n_i b_{ii} - n_i\alpha_i r\sigma_i^2$. Therefore, for this case we have

$$\alpha_i^{DNS} = \frac{n_i b_{ii}}{1+n_i r\sigma_i^2}$$

This gives a level of effort

$$t_i^{DNS} = \frac{\sum_{j=1} n_j b_{ij}}{c_i}\frac{n_i b_{ii}}{\sum_{j=1} n_j b_{ij}}\frac{1}{1+r\sigma_j^2}\frac{1+r\sigma_j^2}{1+n_i r\sigma_i^2} = t_i^{*}\frac{n_i b_{ii}}{\sum_{j=1} n_j b_{ij}}\frac{1}{1+r\sigma_j^2}\frac{1+r\sigma_j^2}{1+n_i r\sigma_i^2} \qquad (9)$$

In this case we have three effects that reduce the level of effort, 1) the externalities, 2) the risk sharing effect and 3) the collective principal problem.

2.3.3 Comparing centralization and decentralization

Comparing (9) with (7) we see that although with centralization there is no problem of externalities, this time it is not clear when the level of effort (and hence welfare) is higher. This is because the agency problem is stronger in the centralized case. The larger the population of principals, the deeper the problem of lack of coordination in contracting with agents. Decentralization will be preferable to centralization whenever the externality effect is less important than the differences of the coordination effect.

$$\frac{n_i b_{ii}}{\sum_{j=1} n_j b_{ij}}\frac{1+r\sigma_j^2}{1+n_i r\sigma_i^2} > \frac{1+r\sigma_j^2}{1+Nr\sigma_j^2} \quad () \quad \frac{n_i b_{ii}}{\sum_{j=1} n_j b_{ij}} > \frac{1+n_i r\sigma_i^2}{1+Nr\sigma_j^2} \qquad (10)$$

To simplify the comparisons, we assume from now on a symmetric case in which $b_{jj} = b$ and $b_{ij} = \phi b$, with $\phi \in [0;1]$. In this case (10) becomes

$$\frac{n_i b}{n_i b + (N - n_i)\phi b} = \frac{n_i}{n_i + (N - n_i)\phi} > \frac{1+n_i r\sigma_i^2}{1+Nr\sigma_j^2}.$$

It is easy to see that

When $\rho = 0$ (no externalities), decentralization is the preferred institutional arrangement; and when $\rho = 1$ (pure public goods), centralization is the preferred institutional arrangement.

More generally, since we know that $\frac{\partial t_i^{DNS}}{\partial \rho} = 0$ and $\frac{\partial t_i^{CNS}}{\partial \rho} = (N - n_i)\frac{b}{(1+N\tau c_j \sigma_j^2)c_j} > 0$, there will be a cut-off point ($\bar{\rho}_j$) such that when $\rho > \bar{\rho}_j$ centralization is preferable and when $\rho < \bar{\rho}_j$, decentralization is preferable..

To find $\bar{\rho}_j$, we have to make $t_i^{DNS} = t_i^{CNS}$. This implies

$$\frac{n_i \tau c_i \sigma_i^2}{1 + n_i \tau c_i \sigma_i^2} = \bar{\rho}_j .$$

Letting

$$\pi_j = n_j \tau C \sigma_j^2$$

we have

$$\frac{\pi_j}{1 + \pi_j} = \bar{\rho}_j$$

which implies that for each π there is a critical ρ above which the centralized solution is better, as shown in Figure 1. π is a measure of the control problems, which are increasing in the variance of the wedge between effort and outcomes σ_j^2, in the conflict of interest between the principals and the agent (the cost c_j) and in the number of principals.

Notice that our result that there are cases in which decentralization is preferable obtains even in the case in which there is homogeneity of preferences across towns.

Since we are assuming that ρ is independent of the region, but the π_j can differ, it could be the case that some goods are better provided by a centralized agent while others by a decentralized one.

2.3.4 Alternative (intermediate) federal arrangements: Regions

The model has compared two situations: one of complete centralization with one of complete decentralization; but perhaps the optimal institutional technology is an intermediate one, call it regionalization. We will see the result when the economies are symmetric and asymmetric.

In a symmetric world ($b_{jj} = b$; $b_{ij} = \theta b$, with $\theta \in [0;1]$ and $n_i = n$ $\forall i$) we can compute the level of effort as a function of the quantity of towns $h \in [1;M]$ that belong to each region. ($h = 1$ would stand for the case of complete decentralization and $h = M$ for the case of complete centralization). From (7),

$$t_j(h) = \frac{\sum_{i=1}^{M} n b_{ij}}{c_j(1 + hnr\, c_j^{3/2})} = \frac{nb + n(h-1)\theta b}{c_j(1 + hnr\, c_j^{3/2})}.$$

It is easy to verify that the sign of $\frac{d\, t_j(h)}{dh}$ is independent of h.[12] Therefore, there is always a corner solution; when the derivative is positive, centralization ($h = K$) is optimal, and when it is negative, decentralization ($h = 1$) is optimal.

The result above depends crucially on the symmetry assumption. We provide now an example of an "asymmetric" country where the optimal institutional technology is neither complete centralization nor complete decentralization.

Imagine there are four towns $i = \{1;2;3;4\}$, with externalities as follows. The commodity of towns 1 and 2 are enjoyed equally by the citizens of towns 1 and 2 ($b_{11} = b_{21}; b_{22} = b_{12}$) and the same happens with the commodities of towns 3 and 4 and its respective citizens ($b_{33} = b_{43}; b_{44} = b_{34}$); while neither the commodities of towns 1 or 2 provide any externality to the citizens of towns 3 and 4 ($b_{13} = b_{14} = b_{23} = b_{24} = 0$), nor the commodities of towns 3 or 4 to the citizens of towns 1 and 2 ($b_{31} = b_{41} = b_{32} = b_{42} = 0$). Using the facts that when there are no externalities, decentralization is the preferred institutional arrangement; and when there is pure public goods, centralization is the preferred institutional arrangement, it is easy to see that the optimal institutional technology is having two regions $\{1;2\}$ and $\{3;4\}$.

[12] Although h is a discrete variable, hence the derivative is not defined, we will overlook this aspect since when the sign of the "derivative" of $t_j(h)$ is independent of the level of h so is the sign of the value of any discrete difference, which guarantees our result.

3 A Recap: On Modelling Decentralization

I

We think that our model provides a useful step in the process of formalizing some of the key concepts being discussed in the decentralization debate around the globe. We provide below a listing of some of the usual claims being heard in favor of the decentralization of political power and public services (see, for instance World Bank, 1999), and try to interpret those claims in more formal language. The "catch-all" expression behind most of those claims is the notion of accountability.

The first channel through which smaller jurisdictions seem to improve political control is the standard Olsonian relationship between group size and free-riding in the voluntary provision of a public good. The benefits of doing independent control are not only diluted by the large number of people sharing the returns but also by the small probability of altering the final output.[13] The application of that logic to the public good of political control is what, in a particular way, we have modelled here. Later on we discuss the generality of this result.

The second oft-mentioned channel is what we might call "the proximity effect." Namely, local officials can be held accountable because they are closer (Ostrom, Schroeder and Wynne, 1993). We interpret this effect as deriving from the fact that citizens and politicians in small communities do interact repeatedly in multiple settings, hence giving the principals (citizens) additional instruments to punish misbehavior in related games - for instance, socially ostracizing a bad governor. (We will argue later that this proximity might also empower local officials to abuse citizens).[14]

[13] Putterman (1993 a,b) uses this logic to study the problem of public ownership.

[14] Notice that this sort of proximity argument might also provide a microfoundation for the association of smaller numbers of people with larger provision of the public good of control (in that case, the emphasis will be on the horizontal relation among principals while in the text we emphasize the relation of each principal to the agent). It is worth reminding that there are some conflicts between these two dimensions since, as

A third channel is that of yardstick competition. Given the standard assumption of unobservable effort, citizens have to infer the governor's behavior from outcomes. If the shocks that create the wedge between effort and outcomes are correlated across jurisdictions, citizens might condition their payments also on outcomes in the other jurisdictions (as in Besley and Case 1995). We conjecture that such extension of our model might generate an increase in the desirability of decentralization.

Another channel might operate through the experimentation/learning possibilities of having multiple jurisdictions. This argument is somewhat tied to some of the previous (or other political-economy) channels, since in principle a centralized government can also experiment over the territory.[15;16]

II

Focusing now on the channel which we have chosen to emphasize, the first one, the size effect, several caveats are in order. First, the intuition that "larger groups will provide smaller amounts of a public good" is not a universal result neither theoretically (for instance, Chamberlin 1974), nor empirically (Isaac and Walker 1988).

This leads to a second point: aggregation technologies (i.e., the way in which individual contributions map into aggregate and individual benefits) do matter, and the incentives resulting from different institutional settings vary according to the nature of the (public) good in question. (See, for instance, the recent paper by Arce and Sandler 1999).

higlighted by our model, each principal might have the incentive of offering a "private" contract.

[15] Regarding the last two channels, one might wonder why do they apply to regions within a country and not across countries. Presumably this is due to the presence of better control variables (in the econometric sense) that allow observers to obtain better information by making intra-county comparisons.

[16] To the previous four arguments one might add, and it is indeed done (World Bank, 1999) the standard Tiebout (1956) argument that when the population is mobile and citizens can "vote with their feet", decentralization may also result in local governments competing with each other to better satisfy the wishes of citizens. As Seabright (1996) forcefully argues, there are conceptual problems in extrapolating the Tiebout results to the centralization/decentralization discussion.

More specifically, of a wide space of possible aggregation technologies, some (but not all) of them will be applicable to the specific problem of principals controlling agents. There are in turn, several possible "technologies" for such control. The particular one we have chosen, is the common-agency model of Bernheim and Whinston (1986) and Dixit (1996).

Even though the common-agency (or multiprincipal) model is a standard one in the literature and did allow us to obtain some insights into the centralization-decentralization question, it is not the most natural framework to think about political control. The archetypical political control technology, voting, is far more restrictive than the set of contracts we have allowed here. In particular, the agent there signs a contract with the whole population, while in our set up, it does so with each citizen.[17] One intriguing possibility would be to explore whether an "optimal" constitutional restriction on the set of contracts that citizens can offer to politicians can lead from the space of contracts we model here to the ones observed in reality. A next step in our agenda will be to embed the decentralization discussion in more explicit political control technologies, and to relate that to this presumably more general framework.

It is clear that, on top of the vertical control mechanism of (retrospective) voting, there are also constitutional arrangements such as division of powers that might also lead to increased government accountability (Persson, Roland and Tabellini 1997). This opens up the door to the modelling of multiprincipal-multiagent situations, which characterize real politics, and to the need of looking into some of the details of more complex governance structures, including the possibility of multiple layers of government operating simultaneously, unlike in our model. The simultaneous presence of various levels of government also requires dealing with multiplicity of public goods (or tasks), something that we have not done here, but can in principle be handled within the common agency framework (as in Dixit's 1996 multitask-multiprincipal model).

[17] Barro (1973) and Ferejohn (1986) are the classics in the economic modeling of principal-agent control through voting. Seabright (1996) takes some elements of Ferejohn's model into the decentralization discussion.

Finally, it is worth pointing out that there are other instruments through which citizens (or groups of citizens) can punish or reward government officials, such as lobbying, campaign contributions, picketing, striking, violence, and other political technologies. Most of these technologies seem to be asymmetrically distributed across citizens, a force that might be behind the "agency rents" we model in a simplistic way here – low effort might be read as policies that favor specific influential groups rather than the general population. Those additional control technologies may also be differentially available in large versus small communities.

III

This leads to another point we want to raise in this section: the drawbacks of decentralization (World Bank 1999). We can organize these drawbacks into those that can be quite directly related to principal-agent problems and those that cannot. We begin with the latter ones.

The most common caveat that enthusiasts of decentralization have in the developing world is the fact that many subnational governments seem to "lack the technical capacities" necessary to undertake many of the decentralized duties. Furthermore, those capacities seem to be unevenly distributed across sub-national units, generating the possibility of a dynamic effect of increasing inequality (imagine for instance the dynamic effect of differential qualities of public education).[18]

[18] The point seems to be a realistic one, but we have some trouble conceptualizing it theoretically. The total pool of human capital is, in principle, independent of the political organization of the country. Why is it the case that any "capacity" available in the centralized case cannot be replicated in the decentralized case? The answer may relate to economies of scale, agglomeration externalities, the fact that smart people do not want to live away of the largest urban centers which have the better amenities, etc. It might be also a transitional effect due to lack of previous experience, or it might relate to more fundamental political economy or institutional issues that do not give local governments the incentives or opportunities to build those capacities. This is a question worth pursuing.

Coming to principal-agent related problems, there are two related caveats about decentralization in the developing world: the risk of \capture" by local elites, and the Madisonian problem of reverse control. Some of our colleagues in Political Science were pretty appalled when they saw us applying a principal (citizen) - agent (governor) framework to think about the possible effects of decentralization in, say, Latin America. Their concerns might be translated as a version of Madison's Dilemma (Kiewiet and McCubbins, 1991). This is a general problem in all agency relations: the resources and authority turned over to the agent for the purpose of furthering the interests of the principals can be turned against the principals. That general agency problem is of particular importance when, as in our case, the agents involved are those in a position of power.

One might speculate that some of those "reverse" control instruments might be more pervasive in smaller communities, perhaps due to reduced political competition within the locality (the downside of the yardstick story) if there is a …xed national pool of political contestants. The increased control of politicians over citizens might also be the downside of the proximity story. It seems promising to attempt to formalize some of these issues within a principal-agent framework.[19]

4 Concluding remarks

We analyze the advantages and disadvantages of centralization in a model with homogeneous people. We …nd that when there are coordination problems among citizens in controlling the government, decentralized political structures could be optimal.

We have focused on the e¢ ciency aspects of the problem. Some of the solutions found are consistent with many different distributions, the distributive aspects jointly with the political arrangements will determine whether the e¢ cient organization will be reached or not, it is not hard to imagine situations where e¢ cient outcomes are dominated politically

[19] Perhaps one way of getting at that problem might be by changing the standard assumption aboout the bargaining technology in principal-agent models.

by suboptimal ones.

As already mentioned, the common agency framework does not fully capture the problem of political control by citizens. The framework assumes that each citizen signs a contract with the agent, while in reality some of these "contracts" are signed collectively through the aggregation of some actions of principals such as voting. (Moreover in many cases it is prohibited that a member of the population signs a contract with the agent to act on his behalf.) This reinforces the claim of the need of study a broader class of "collective" principal problem.

Even though the "generic" agency model we have chosen has limitations to study political applications (as those listed above), it has the advantage of allowing us to link with other areas of application. For instance, our results could be of some use in the theory of the firm: for instance the coordination necessary for agency control will influence the optimal ownership structure of firms, the optimal size and configuration of the firms and therefore might affect market structures.[20]

Finally, we have not yet fully exploited the framework in order to answer the fundamental question of exactly what goods, under what circumstances will be provided by different levels of government. We can give some partial answers by varying some of the parameters (such as b_{ij}) in our model, but there are types of public goods not captured by our production/consumption technology. Furthermore, we also need to look at a multi-good economy. Several of these steps can be taken from the framework we used here, and constitute the next steps in the agenda.

[20] Our problem is similar to the problem of controlling the managers of a firm with disperse ownership. Schleifer and Vishny (1986) propose having one big shareholder with very strong incentives to control the agent as a solution to that problem. It seems hard to apply such a solution to our multi-layer government case; we cannot give to a citizen neither the incentives nor the right to make him behave as a big shareholder.

References

Arce D. and Sandler (1999) "Regional Public Goods: Strategies and Institutions" Mimeo.

Bardhan P. and Mookherjee (1998) "Expenditure Decentralization and the Delivery of Public Services in Developing Countries" Mimeo.

Barro,-Robert-J (1973) "The Control of Politicians: An Economic Model" Public-Choice; 14, Spr. pages 19-42.

Bernheim D. and Whinston, M.D. (1986) "Common Agency" Econometrica Vol.54, N 4, 923-942.

Besley, T. and Case, A. (1995) "Incumbent Behavior: Vote Seeking, Tax Setting and Yardstick Competition" American Economic Review, 85, 1, 25-45.

Besley and Coate (1998) "Centralized versus Decentralized Provision of Local Public Goods: A Political Economy Analysis" mimeo.

Chamberlin John (1974) "Provision of Public Goods as a function of Group size" American Political Science Review LXVIII 707-16.

Dixit, Avinash (1996) "Multiprincipal Multitask Agency", appendix to The Making of Economic Policy. A Transaction-Cost Politics Perspective. MIT Press.

Dixit, A., G. Grossman and E. Helpman (1997) "Common Agency and Coordination: General Theory and Application to Government Policy Making" Journal-of-Political-Economy; 105(4), pages 752-769.

Ferejohn,-John (1986) "Incumbent Performance and Electoral Control" Public-Choice; 50 (1-3), pages 5-25.Reprinted in Persson,-Torsten; Tabellini,-Guido, eds. Monetary and …scal policy. Volume 2. Politics. Cambridge and London: MIT Press, pages 29-45.

Grossman,-Gene-M.; Helpman,-Elhanan (1994)"Protection for Sale" American-Economic-Review; 84(4), September, pages 833-50.

Grossman,-Gene-M.; Helpman,-Elhanan (1995) "The Politics of Free-Trade Agreements" American-Economic-Review; 85(4), September, pages 667-90.

Holmstrom, Bengt and Paul Milgrom (1987) "Aggregation and Linearity in the Provision of Intertemporal Incentives" Econometrica 55, nº 2: 303-328.

Holmstrom, Bengt and Paul Milgrom (1991) "Multitask Principal Agent Analysis: Incentive Contracts, Asset ownership, and Job Design" Journal of Law, Economics and Organization 7, Special Issue: 24-51.

Isaac, R.M. and Walker, J.M. (1988) "Group size effects in public goods provision: The voluntary contributions Mechanism" The Quarterly Journal of Economics, Vol. 103, Issue 1, 179-199.

Kieweit and McCubbins (1991) The Logic of Delegation University of Chicago Press.

Lockwood, Ben (1998) "Distributive Politics and the Benefits of Decentralization" CSGR Working Paper N 10/98.

Oates, Wallace (1972) Fiscal Federalism. Harcourt Brace: New York.

Olson, M Jr. (1968) The Logic of Collective Action: Public Goods and the Theory of Groups. New York: Schocken

Ostrom, Schroeder and Wynne (1993) Institutional Incentives and Sustainable Development. Boulder: Westview Press.

Persson, Roland and Tabellini (1997) "Separation of Powers and Political Accountability" Quarterly Journal of Economics November, 1163-1202.

Persson,-Torsten; Tabellini,-Guido (1996) "Federal Fiscal Constitutions: Risk Sharing and Moral Hazard" Econometrica; 64(3), May, pages 623-46.

Persson,-Torsten; Tabellini,-Guido Federal (1996b) "Fiscal Constitutions: Risk Sharing and Redistribution" Journal-of-Political-Economy; 104(5), October, pages 979-1009.

Putterman, Louis (1993 a) "Exit, Voice, and Portfolio Choice: Agency and Public Ownership," Economics and Politics 5: 205-18, 1993.

Putterman, Louis (1993 b) "Incentive Problems Favoring Noncentralized Investment Fund Ownership," pp.156-68 in P. Bardhan and J.E. Roemer, eds., Market Socialism: The Current Debate. New York: Oxford University Press.

Qian and Weingast (1997) "Federalism as a Commitment to Preserving Market Incentives" Journal of Economic Perspectives Fall, 83-92.

Rodrik,-Dani (1995) "Political Economy of Trade Policy" in Grossman,-Gene-M.; Rogoff,-Kenneth, eds. Handbook of International Economics. Volume 3. Handbooks in Economics. Amsterdam; New York and Oxford: Elsevier, North-Holland, pages 1457-94.

Seabright, Paul (1996) "Accountability and decentralization in government: An incomplete contracts model" European Economic Review 40, 61-89

Schleifer and Vishny (1986) "Large Shareholders and Corporate Control" Journal of Political Economy, vol96, nro31, 461-488.

Schleifer and Vishny (1997) "A Survey of Corporate Governance" The Journal of Finance, vol 52, nro 2, 737-783.

Stole, Lars A. (1997) "Mechanism Design under Common Agency: Theory and Applications" Mimeo.

Tiebout, Charles (1956) "A Pure Theory of Local Expenditures" Journal-of-Political-Economy; 64(5), pages 416-424.

World Bank (1999) (Burki, Perry and Dillinger) Beyond the Center. Decentralizing the State. Washington, D.C.