Centralization vs. Decentralization: A Principal-Agent Analysis

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

Abstract

The architecture of public decision making is changing throughout the world through processes of economic integration and of decentralization. The decision to allocate policy jurisdictions to different levels of government is related to a number of trade-offs between the advantages and disadvantages of centralized versus decentralized provision. A trade-off central to many discussions is that between the internalization of externalities under centralization versus an "accountability" advantage of decentralization. In this paper we formalize this trade-off in 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]A previous version of this paper has circulated with the title "A Principal-Agent Building Block for the Study of Decentralization and Integration". We thank the research assistance of Gabriel Basaluzzo, and the valuable comments of Steve Coate, Ernesto DalBo, Dhammika Dharmapala, Avinash Dixit, Jeff Frieden, Alex Galetovic, Matías Iaryczower, Ines Macho-Stadler, Pablo Spiller, Ernesto Stein, Andrés Velasco, Marcelo Veracierto and seminar participants at Universidad de San Andrés, UTDT, CEMA, NYU, at the Center of Studies for Basic Research in the Social Sciences (Harvard University), at the Meeting of the Latin America and the Caribbean Economic Association in Santiago de Chile, at the Annual Conference of the Central Bank of Uruguay, at the Political Economy Group of LACEA meeting in Cartagena, and at PET 02-Third International Conference on Public Economics.

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). Both processes are the two faces of the same coin: the search for appropriate governance structures for public decision making.

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 cast 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 coordination 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.

We formalize the trade off between one of the main advantages of centralized decision making - namely, the internalization of externalities - and one of its main disadvantages - namely the "democratic deficit" of having decision making further removed from the citizenry. In particular, we cast the latter dimension in a principal-agent framework.[2]

We analyze a case in which the principal is not a single individual but a group, a population. There are, in principle, several stages of an agency problem where one can introduce a collective control problem: the contract stage, the monitoring stage, the enforcement stage. 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 class of models that has become the workhorse multi principal-agent framework: the "common agency" model (Bernheim and Whinston 1986, Grossman and Helpman 1994, Dixit 1996), which focuses on the contracting stage. One

[2]In a sense, we consider an exercise in "optimal constitutional choice" between centralization and decentralization, recognizing (in a stylized manner) some political economy issues within each regime. In the concluding section we briefly reflect on the politics of such "constitutional" choice.

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.[3]

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, under some conditions, 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. (Oates, 1999).

Later work has emphasized, hence, that the case for decentralization has to be driven by political economy considerations. Besley and Coate (1998), Lockwood (2002) 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; in their model a central government has limited ability to monitor the performance of the bureaucrats while in a de-

[3] Ours is a model of representative (not direct) democracy. We assign the policymaker/s the right to choose policy, and we give the citizens a "vainilla" principal-agent control mechanism. There are some interesting papers by Persson and Tabellini (1996 and 1996b), Cremer and Palfrey (1999), and others looking at direct-democracy political aggregation technologies. Those models derive rich implications, but by construction ignore issues of political agency.

centralized system the local governments may be subject to capture by local elites. Besharov (2001) studies different regimes for the provision of local public goods in a "menu auction" common agency setting. In his model, the advantage of the decentralized regime is that it reduces influence costs.

Many of the papers in this literature require interjurisdictional heterogeneity "a la Oates" in order to derive benefits of decentralization. (Besley and Coate 1998 and Seabright 1996 are exceptions.) 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. We also discuss the limitations of this rather standard principal-agent setup to capture some key elements of the decentralization debate.

Section 2 presents the model and the analysis. In Section 3 we cast the recent decentralization discussion in more theoretical terms; and we assess how far does the standard multiprincipal agent model travel in addressing some of those applied concerns. Section 4 concludes.

2 The model[4]

There are M towns. A "local public good" has to be provided for each town. (This could be a metaphor for more general policies that do have asymmetric regional effects.) Hence, we have an M goods economy $\mathbf{x} = (x_1, x_2, ..., x_M)$. There are $N = n_1 + n_2 + ... + n_M$ citizens (principals) of type $1, 2, ...M$ respectively.

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

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

$$b_{i1} \cdot x_1 + b_{i2} \cdot x_2 + ... + b_{iM} \cdot x_M = \mathbf{b}_i' \cdot \mathbf{x} \tag{1}$$

$b_{ii} \geq 0$ is the utility that each principal of type i gets for a unit of his own local public good and $b_{ij} \geq 0$ $(i \neq j)$ 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 second case, "*decentralization,*" we do not allow contracting between citizens in one locality and policymakers in another.

The production technology in each locality is given by a level of "effort" (t_i) chosen by the agent responsible to provide the local public good in that town plus an error term (ε_i). The error terms are independently and normally distributed with mean 0 and variance σ_i^2. [5] The output vector is $\mathbf{x} = \mathbf{t} + \boldsymbol{\varepsilon}$,where $\mathbf{t}$ is the vector of the agent(s)' efforts, $\mathbf{t} = (t_1, t_2, ..., t_n)$, and $\boldsymbol{\varepsilon} \in R^M$ 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 $\frac{1}{2}\mathbf{t}'C\mathbf{t}$ where[6]

$$C = \begin{bmatrix} c_1 & 0 & 0 & \cdots & 0 \\ 0 & c_2 & 0 & \cdots & 0 \\ 0 & 0 & c_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & c_M \end{bmatrix}. \tag{2}$$

[5] In a more general case there will be a variance matrix which might include non-zero off-diagonal elements. In that case, the correlation of shocks will allow local citizens to condition rewards on comparative performance (i.e., yardstick competition).

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

Hence when there is only one agent, his "monetary" payoff is

$$w = z - \frac{1}{2}\mathbf{t}'C\mathbf{t} = z - \frac{1}{2}\sum_{j=1}^{M} c_j t_j^2; \tag{3}$$

and when there are M agents, their payoffs are

$$w_j = z_j - \frac{1}{2}t_j^2 c_j. \tag{4}$$

The expected utility of principal i is $\sum_{j=1}^{M} b_{ij}t_j - z_i$. The expected utility of the "aggregate" principal is[7]

$$\sum_{j=1}^{M}\left(\sum_{i=1}^{M} n_i b_{ij}\right) t_j - z, \tag{5}$$

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

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 the principals act as unified actors – there is no problem of cooperation among principals in contracting with the agent. In 2.1.1 effort is observable and verifiable (hence contractable), while in 2.1.2 effort is not observable. In subsection 2.2 effort is not observable and principals (within political jurisdictions) act in a non-coordinated

[7]We use the notation j to refer to goods, and i to refer to principals' type, although in the solutions we will use them interchangeably, since sometimes (centralized case) we will emphasize the agent's choice of effort in dimension j, and other times (decentralized case) we will focus on the incentive scheme provided by principals of type i.

manner.[8] We will use the notation t_j^{1c} to denote the level of effort in producing j under a centralized political structure in case 1 (contractable effort and united principals), t_j^{3d} to denote the level of effort in producing j under a decentralized political structure in case 3 (non-contractable effort and separate principals), and so forth.[9]

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 2.2.) The timing is the standard one: the contract/s is/are offered by the principals, the agent/s accept or reject (leading to the participation constraint), agent supplies effort (leading to the incentive compatibility constraint), shocks are realized, and then outcomes and payoffs obtained.

2.1 A benchmark: United principals

2.1.1 Contractable effort

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

Centralized case (first-best)

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

[8] We follow Dixit (1996) and call these cases *united* and *separate* principals, respectively. In 2.1 principals (whithin a political jurisdiction) coordinate fully. In section 2.2, they do not coordinate at all. (In each case, we compare a unique national political jurisdiction, to multiple local jurisdictions). In the concluding section we speculate about intermediate degrees of cooperation, possible institutional arrangements for that, and how that might be affected by multi-layered political jurisdictions.

[9] We do not analyze explicitly a fourth possible context with contractable effort and separate principals. Results analogous to the case of contractable effort and united principals can be obtained in that case (see Bernheim and Whinston, 1986).

surplus,

$$\sum_{j=1}^{M}\left[\sum_{i=1}^{M} n_i b_{ij} t_j - \frac{1}{2} c_j t_j^2\right] \tag{6}$$

The first order condition with respect to t_j, leads to

$$\sum_{i=1}^{M} n_i b_{ij} = c_j t_j. \tag{7}$$

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

$$t_j^{1c} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_j} = t_j^* \tag{8}$$

for all j, which corresponds to the first best level t_j^*.

We will use this case 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.[10]

Decentralized case

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

As before, the payment from principal to agent (now, in each locality) is just a transfer that leaves each agent in its reservation utility level. The aggregate principal of each locality will, thus, choose the level of effort that maximizes the aggregate surplus in 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 $n_i b_{ii} = c_i t_i$. They equate marginal social cost to marginal social benefit *of the locality*. Although effort is contractable the result is not optimal since each principal does

[10] 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 (6) to capture that loss. We show later that t is also a sufficient statistic for welfare in that case.

not take into account the externalities that its good provides to the other regions. The level of effort in this decentralized world is

$$t_i^{1d} = \frac{n_i b_{ii}}{c_i} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_i} \frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}} = t_i^* \frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}}. \tag{9}$$

There is under-provision of effort, which comes from a standard Nash externality across localities. Therefore, with contractable effort, centralization is the preferable institutional arrangement.

2.1.2 Non contractable effort

Centralized case (the standard "second-best" principal-agent case)

Now principals can monitor the efforts $\mathbf{t}$ but only imperfectly, i.e., they observe $\mathbf{x}$ but not $\mathbf{t}$. The agent is offered the contract $z = \boldsymbol{\alpha}'\mathbf{x} + \beta$. 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.) $\boldsymbol{\alpha}$ 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 $\boldsymbol{\alpha}'\mathbf{t} + \beta - \frac{1}{2} r \boldsymbol{\alpha}' \Omega \boldsymbol{\alpha} - \frac{1}{2} \mathbf{t}' C \mathbf{t}$, and the principals' benefit is $\sum_{j=1}^{M} \left(\sum_{i=1}^{M} n_i b_{ij} - \alpha_j \right) x_j - \beta$.

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

$$\max_{\mathbf{t}} \quad \sum_{j=1}^{M} \alpha_j t_j - \frac{r}{2} \sum_{j=1}^{M} \alpha_j^2 \sigma_j^2 - \frac{1}{2} \sum_{j=1}^{M} c_j t_j^2 + \beta \tag{10}$$

which leads to

$$t_j = \frac{\alpha_j}{c_j} \tag{11}$$

for all j.

The expected social surplus is then

$$\sum_{j=1}^{M} \left[\left(\sum_{i=1}^{M} n_i b_{ij} \right) \frac{\alpha_j}{c_j} - \frac{r}{2} \alpha_j^2 \sigma_j^2 - \frac{1}{2} c_j \left(\frac{\alpha_j}{c_j} \right)^2 \right] \tag{12}$$

The principal maximizes (12) with respect to $\boldsymbol{\alpha}$ leading to $\alpha_j^{2c} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{\left(1+rc_j\sigma_j^2\right)}$, which implies

$$t_j^{2c} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_j} \frac{1}{\left(1+rc_j\sigma_j^2\right)} = t_j^* \frac{1}{\left(1+rc_j\sigma_j^2\right)} \tag{13}$$

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 $\alpha_j < \alpha_j^{2c}$, and so is t_j since it is increasing in α_j when $t_j < t_j^{2c}$. Hence, as in the case when effort was contractable, t is a measure of welfare.

Decentralized case

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

$$\sum_{j=1}^{M} n_i b_{ij} \frac{\alpha_j}{c_j} - \frac{r}{2}\alpha_i^2\sigma_i^2 - \frac{1}{2}c_i\left(\frac{\alpha_i}{c_i}\right)^2 \tag{14}$$

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

$$t_i^{2d} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_i} \frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}} \frac{1}{(1+c_i r\sigma_i^2)} = t_i^* \frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}} \frac{1}{(1+c_i r\sigma_i^2)} \tag{15}$$

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 (15), in this case there are two sources of the under provision of effort, the externalities ($n_i b_{ii} < \sum_{i=1}^{M} n_i b_{ij}$) 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, a harder agency problem due to the larger number of principals. To that we turn now.

2.2 Non contractable 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) in their focus on the contracting stage of the control problem.[11] (Other stages of control, such as monitoring or, perhaps, enforcement can be also subject to the type of 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.[12] We can imagine two modelling choices at that point: either the agent

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

[12] 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. 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. It also turns out that the second stage, not always made explicit in the literature, might also impinge upon the "bargaining" outcome. (See Tommasi and Weinschelbaum, 2001).

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*.[13] An example of intrinsic common agency would be regulation by multiple authorities: the regulated firm can choose between abiding to all of those payment functions or exiting the market altogether. An example of "delegated" common agency would be a salesman, distributor or retailer who decides whose products to carry. Most 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.[14]

We consider below our two standard institutional comparisons. In the *centralized* case, each of the N total principals will offer a contract to the only agent. Each contract might include the realized level of output in all the activities. In a Nash equilibrium, each principal offers his contract, taking all the other principal's contracts as given.

In the *decentralized* case each principal in town i is allowed to offer a contract to the agent of that town but not to agents from other towns. Within each town, each principal offers his contract, taking all the other principal's contracts as given. One way of thinking

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

[14]For useful steps at modelling delegated common agency, see Bernhein and Whinston (1985), Bisin and Guaitoli (1999) and Kahn and Mookherjee (1998).

about our exercise would be to imagine that in the *centralized* case there are still M agents, but that "cross-contracts" are allowed.

2.2.1 Centralized case

In this case each of the $N = \sum_{i=1}^{M} n_i$ principals can contract with the one agent. Principals can monitor efforts $\mathbf{t}$ only imperfectly. The agent is hired through individual contracts with each principal l and payment $z^l = \boldsymbol{\alpha}^{l\prime}\mathbf{x} + \beta^l$, where $\boldsymbol{\alpha}^l = (\alpha_1^l, \alpha_2^l, ..., \alpha_M^l)$. Each principal will offer a contract, taking as given the contracts offered by the other principals. Let $\alpha_j = \sum_l \alpha_j^l$ and $\beta = \sum_l \beta^l$. The agent's CE is $\boldsymbol{\alpha}'\mathbf{t} + \beta - \frac{r}{2}\boldsymbol{\alpha}'\Omega\boldsymbol{\alpha} - \frac{1}{2}\mathbf{t}'C\mathbf{t}$.

At the stage of choosing effort, the agent maximizes

$$\underset{\mathbf{t}}{max} \qquad \sum_{j=1}^{M} \alpha_j t_j - \frac{r}{2}\sum_{j=1}^{M} \alpha_j^2\sigma_j^2 - \frac{1}{2}\sum_{j=1}^{M} c_j t_j^2 + \beta \tag{16}$$

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

$$\sum_{j=1}^{M}\left[\frac{\left(A_j^l\right)^2}{c_j} - \frac{r}{2}\left(A_j^l\right)^2\sigma_j^2 - \frac{1}{2}\frac{\left(A_j^l\right)^2}{c_j}\right] + B^l = \sum_{j=1}^{M}\left(A_j^l\right)^2\left[\frac{1}{2c_j} - \frac{r}{2}\sigma_j^2\right] + B^l \tag{17}$$

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

$$\sum_{j=1}^{K}\left(A_j^l + \alpha_j^l\right)^2\left[\frac{1}{2c_j} - \frac{r}{2}\sigma_j^2\right] + B^l + \beta^l. \tag{18}$$

The marginal gains in signing the contract are

$$\sum_{j=1}^{M}\left(\left(\alpha_j^l\right)^2 + 2A_j^l\alpha_j^l\right)\left(\frac{1}{2c_j} - \frac{r}{2}\sigma_j^2\right) + \beta^l. \tag{19}$$

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

$$\sum_{j=1}^{M}\left[b_{ij}\frac{\alpha_j^l}{c_j}-\alpha_j^l\frac{A_j^l+\alpha_j^l}{c_j}-\beta^l\right]. \tag{20}$$

From (19) and (20), the total bilateral surplus is then

$$\sum_{j=1}^{M}\left[b_{ij}\frac{\alpha_j^l}{c_j}-\alpha_j^l\frac{A_j^l+\alpha_j^l}{c_j}+\left(\left(\alpha_j^l\right)^2+2A_j^l\alpha_j^l\right)\left(\frac{1}{2c_j}-\frac{r}{2}\sigma_j^2\right)\right]. \tag{21}$$

Maximizing (21)with respect to α_j^l leads to

$$\begin{aligned} 0 &= \frac{b_{ij}}{c_j}-\frac{A_j^l+2\alpha_j^l}{c_j}+\left(\alpha_j^l+A_j^l\right)\left(\frac{1}{c_j}-r\sigma_j^2\right)= \\ &= b_{ij}-\alpha_j^l-\left(\alpha_j^l+A_j^l\right)c_j r\sigma_j^2. \end{aligned}$$

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

$$\alpha_j=\sum_{i=1}^{M}\alpha_j^l=\sum_{i=1}^{M}n_i b_{ij}-N\alpha_j rc_j\sigma_j^2. \tag{22}$$

Therefore, for this case we have

$$\alpha_j^{3c}=\frac{\sum_{i=1}^{M}n_i b_{ij}}{1+Nrc_j\sigma_j^2}. \tag{23}$$

This gives a level of effort

$$t_j^{3c}=\frac{\sum_{i=1}^{M}n_i b_{ij}}{c_j}\frac{1}{1+rc_j\sigma_j^2}\frac{1+rc_j\sigma_j^2}{1+Nrc_j\sigma_j^2}=t_j^*\frac{1}{1+rc_j\sigma_j^2}\frac{1+rc_j\sigma_j^2}{1+Nrc_j\sigma_j^2}, \tag{24}$$

which is smaller than in the case of united principals. The first term in the right hand side of (24) is the optimal value of t_j but this is multiplied by the risk-sharing effect $\frac{1}{1+rc_j\sigma_j^2}$, and by the "too many principals" effect $\frac{1+rc_j\sigma_j^2}{1+Nrc_j\sigma_j^2}$.

2.2.2 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 l of his town is $\left(\left(\alpha_i^l\right)^2 + 2A_i^l\alpha_i^l\right)\left(\frac{1}{2c_i} - \frac{r}{2}\sigma_i^2\right) + \beta^l$.

Principal l's expected utility in the absence of this relationship is $\sum_{j=1}^{M} b_{ij}\frac{A_j^l}{c_j}$, whereas if he signs the contract it will be $\sum_{j=1}^{M} b_{ij}\frac{A_j^l+\alpha_j^l}{c_j} - \alpha_i^l\frac{A_i^l+\alpha_i^l}{c_i} - \beta^l$. The total bilateral surplus is then

$$\sum_{j=1}^{M} b_{ij}\frac{A_j^l+\alpha_j^l}{c_j} - \alpha_i^l\frac{A_i^l+\alpha_i^l}{c_i} + \left(\left(\alpha_i^l\right)^2 + 2A_i^l\alpha_i^l\right)\left(\frac{1}{2c_i} - \frac{r}{2}\sigma_i^2\right). \tag{25}$$

Maximizing (25) with respect to α_i^l leads to

$$0 = b_{ii} - \alpha_i^l - \left(\alpha_i^l + A_i^l\right) rc_i\sigma_i^2. \tag{26}$$

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

$$\alpha_i^{3d} = \frac{n_i b_{ii}}{1 + n_i rc_i\sigma_i^2} \tag{27}$$

This gives a level of effort

$$t_i^{3d} = \frac{\sum_{i=1}^{M} n_i b_{ij}}{c_i}\frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}}\frac{1}{1+rc_j\sigma_j^2}\frac{1+rc_j\sigma_j^2}{1+n_i rc_i\sigma_i^2} = t_i^*\frac{n_i b_{ii}}{\sum_{i=1}^{M} n_i b_{ij}}\frac{1}{1+rc_j\sigma_j^2}\frac{1+rc_j\sigma_j^2}{1+n_i rc_i\sigma_i^2} \tag{28}$$

In this case we have three forces reducing the level of effort, 1) the externalities, 2) the risk sharing effect and 3) the "many principals" problem.

2.2.3 Comparing centralization and decentralization

Comparing (28) with (24) 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, or

$$\frac{n_j b_{jj}}{\sum_{i=1}^{M} n_i b_{ij}} \frac{1+rc_j\sigma_j^2}{1+n_j rc_j\sigma_j^2} > \frac{1+rc_j\sigma_j^2}{1+Nrc_j\sigma_j^2} \iff \frac{n_j b_{jj}}{\sum_{i=1}^{M} n_i b_{ij}} > \frac{1+n_j rc_j\sigma_j^2}{1+Nrc_j\sigma_j^2}. \tag{29}$$

To simplify the comparisons, we assume from now on a symmetric case in which $b_{jj} = b$ and $b_{ij} = \gamma b$, with $\gamma \in [0,1]$. In this case (29) becomes

$$\frac{n_j b}{n_j b + (N-n_j)\gamma b} = \frac{n_j}{n_j + (N-n_j)\gamma} > \frac{1+n_j rc_j\sigma_j^2}{1+Nrc_j\sigma_j^2}. \tag{30}$$

It is easy to see that:

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

More generally, since we know that $\frac{\partial t_i^{3d}}{\partial \gamma} = 0$ and $\frac{\partial t_i^{3c}}{\partial \gamma} = (N-n_j)\frac{b}{\left(1+Nrc_j\sigma_j^2\right)c_j} > 0$, there will be a cut-off point $(\overline{\gamma}_j)$ such that when $\gamma > \overline{\gamma}_j$ centralization is preferable and when $\gamma < \overline{\gamma}_j$, decentralization is preferable..

To find $\overline{\gamma}_j$, we have to make $t_j^{3d} = t_j^{3c}$. This implies

$$\frac{n_j rc_j\sigma_j^2}{1+n_j rc_j\sigma_j^2} = \overline{\gamma}_j. \tag{31}$$

Letting

$$\delta_j = n_j rc_j\sigma_j^2 \tag{32}$$

we have

$$\frac{\delta_j}{1+\delta_j} = \overline{\gamma}_j \tag{33}$$

which implies that for each δ there is a critical γ above which the centralized solution is better, as shown in Figure 1. δ 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 citizens/principals in the town.

<insert Figure 1>

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. We do not require heterogeneity a la Oates to make a case for decentralization.[15]

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

2.2.4 Alternative (intermediate) federal arrangements: Regions

As an illustration of possible applications of the model to study alternative institutional questions, we consider here intermediate federal arrangements. The model so far 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*, grouping some but not all localities together. It turns out that the answer to that question depends heavily on the degree of heterogeneity and on the diversity of externalities across towns. We consider first a fully (symmetric) world.

In a symmetric world ($n_i = n\ \forall i$,$b_{jj} = b\ \forall j$,and $b_{ij} = \gamma b\ \forall i\ \forall j$, with $\gamma \in [0,1]$) 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 (24),

$$t_j(h) = \frac{\sum_{i=1}^{h} nb_{ij}}{c_j(1 + hnrc_j\sigma_j^2)} = \frac{nb + n(h-1)\gamma b}{c_j(1 + hnrc_j\sigma_j^2)}. \tag{34}$$

[15] A similar result is obtained by the other paper that emphasizes the accountability dimension of decentralization, Seabright (1996).

It is easy to verify that the sign of $\frac{dt_j(h)}{dh}$ is independent of h. 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.[16] (As we already know, the derivative will be positive iff $\frac{\delta_j}{1+\delta_j} < \gamma_j$.)

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 $\{3, 4\}$.

Of course, less extreme examples can be constructed, but the stark case we analyzed captures the flavor of the more general cases: regions should be organized trading off the internalization of externalities with the loss of agency control.

3 A recap: on modelling decentralization

I

We think that our model provides a useful step in the process of formalizing some of the

[16] Of course h is a discrete variable, and hence the derivative is not defined. But 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. This guarantees our result.

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 one citizen controlling the government are not only diluted by the large number of people sharing the returns but also by the small probability of altering the final output.[17] 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.[18] A reader suggested that the proximity-effect might relate to informational issues: i.e. it is easier to observe effort at the local level, which might be captured by σ_j^2 being smaller in local as opposed to national provision.[19]

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

[18] This proximity might also empower local officials to abuse citizens, the so-called Madisonian problem of reverse control. One might speculate that some of those reverse control instruments mught be more pervasive in smaller communities (and specially in developing countries.) It might be promising to attempt to formalize some of these issues within a principal-agent framework. Bardhan and Mookherjee (1999) take some steps in that direction.

[19] 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

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 (localities as "laboratories of democracy.") 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. Looking at policy experimentation under decentralization, Strumpf (1999) makes the perceptive point that since experimentation creates externalities, it might also be subject to Nash problems. But if decentralization does indeed (as Strumpf finds) increase experimentation, it might be due to the increased incentives of local politicians through channels like the ones we emphasize in this paper.[20]

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. Seabright (1996) forcefully argues that there are conceptual problems in extrapolating the Tiebout results to the centralization/decentralization discussion. Also, we agree with Bardhan and Mookherjee (1998)'s point

to the agent). It is worth reminding that there are some conflicts between these two dimensions since, as higlighted by our model, each principal might have the incentive of offering a "private" contract.

[20]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-country comparisons. Similarly, it has been argued in the literature on macroeconomic policy choice, that countries are more likely to imitate the policies of **similar** countries (Tommasi and Velasco, 1996, and Meseguer 1999).

that the assumption of mobility of fully informed citizens in search of a perfect match between their preferences and public services seems less applicable in less developed countries.

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, Arce and Sandler 1999).

More specifically, of a wide space of possible aggregation technologies, some (but not all) 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 might not the most natural framework to think about political control. The archetypal 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.[21] 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

[21] Barro (1973) and Ferejohn (1986) are the classics in the economic modeling of principal-agent control through voting. Personn and Tabellini (2000) give an excellent textbook account of the state of modelling in that front. Seabright (1996) takes some elements of Ferejohn's model into the decentralization discussion.

discussion in more explicit political control technologies, and to relate the results of such analysis to the ones obtained under the presumably more general framework utilized here.

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.[22]

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. (See Bardhan and Mookherjee 1999, for one approach to this problem.)

[22]Political parties are natural "agreggators" of citizens in controlling the government. But that begs the question of how parties are controlled. The formal economic modelling of parties might shed some light on these issues, specially if it takes into consideration the differential national or local centerdness of party power (see Jones, Saiegh, Spiller and Tommasi 2001 and 2002 for a discussion of some of this in the case of Argentina).

4 Concluding remarks

We analyze the advantages and disadvantages of centralization and decentralization, and we find that when there are coordination problems among citizens in controlling the government, decentralized political structures could be optimal, even if all localities have the same preferences.

We focused on "efficiency" aspects of the problem. Some of the solutions found are consistent with many different distributions; the distributive aspects jointly with the pre-existing political arrangements will determine whether the efficient organization will be reached or not. It is not hard to imagine situations where efficient outcomes are dominated politically by suboptimal ones. (Lockwood 1998 provides interesting examples of such situations.)

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.[23]

[23] 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; although we might think of political parties playing a similar role as intermediaries.

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 issues could be addressed building from the framework we used here, and constitute the next steps in the agenda.

References

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

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

Bardhan, P. and Mookherjee, D. (1999) "Relative Capture of Local and Central Governments. An Essay in the Political Economy of Decentralization" Mimeo.

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

Bernheim, D. and Whinston, M.D. (1985) "Common Marketing Agency as a Device for Facilitating Collusion" *Rand Journal of Economics*, 16 , 269-281.

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

Besharov G. (2001) "Influence Costs in the Provision of Local Public Goods" mimeo.

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

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

Bisin, A. and Guaitoli, D. (1999) "Moral Hazard and Non-Exclusive Contracts" Mimeo, NYU and Pompeu Fabra.

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

Cremer, J. and Palfrey, T. (1999) "Political Confederation" *American Political Science Review*, 93(1): 69-83. March.

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

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

Ferejohn, J. (1986) "Incumbent Performance and Electoral Control" *Public Choice*; 50 (1-3), pages 5-25.

Grossman,G. and Helpman,E. (1994) "Protection for Sale" *American Economic Review*; 84(4), September, pages 833-50.

Grossman,G. and Helpman,E. (1995) "The Politics of Free-Trade Agreements" *American Economic Review*; 85(4), September, pages 667-90.

Holmstrom, B. and Milgrom, P. (1987) "Aggregation and Linearity in the Provision of Intertemporal Incentives" *Econometrica* 55, n^o 2: 303-328.

Holmstrom, B. and Milgrom, P. (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.

Jones, M., Saiegh, S., Spiller, P. and Tommasi, M. (2001) "Keeping a Seat in Congress: Provincial Party Bosses and the Survival of Argentine Legislators." Mimeo.

Jones, M., Saiegh, S., Spiller, P. and Tommasi, M. (2002) "Amateur Legislators - Professional Politicians: The Consequences of Party-Centered Electoral Rules in a Federal System" *American Journal of Political Science* July.

Kahn, C. and Mookherjee, D. (1998) "Competition and incentives with nonexclusive contracts" *RAND Journal of Economics*, 29(3), Autumn, pp. 443-465.

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

Meseguer, C. (2000) "Learning and Policy Choices with an Application to IMF Agreements" Mimeo, New York University, March.

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

Oates, W. (1999) "An Essay on Fiscal Federalism" *Journal of Economic Literature*, Vol XXXVII, September, pp. 1120-1149.

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

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

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

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

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

Persson, T. and Tabellini,G. (2000) *Political Economics. Explaining Economic Policy.* MIT Press.

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

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

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

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

Schleifer, A. and Vishny, R. (1986) "Large Shareholders and Corporate Control" *Journal of Political Economy*, vol 96 (31), 461-488.

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

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

Strumpf, K. (1999) "Does Government Decentralization Increase Policy Innovation?" Mimeo, University of North Carolina, Chapel Hill.

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

Tommasi, M. and Velasco, A. (1996) "Where Are We in the Political Economy of Reform?" *Journal of Policy Reform*, Vol 1, April.

Tommasi, M. and Weinschelbaum, F. (2001) "The Threat of Insurance: On the Robustness of Principal-Agent Models" Mimeo, Universidad de San Andrés.

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