A Political Economy Theory of Partial Decentralization*

John William Hatfield
Graduate School of Business
Stanford University

Gerard Padró i Miquel
London School of Economics
NBER and BREAD

July 2008

Abstract

We revisit the classical problem of tax competition in the context of federal nations. We show that when all public goods are provided by the central government, capital taxes may be inefficiently high, as the median voter favors capital taxes over non-distortive lump-sum taxes if there is inequality in capital holdings. As a consequence of high expected capital taxes, capital generation is inefficiently low. The median voter would therefore like to commit to a lower level of capital taxes. Decentralization provides such a commitment: local governments avoid using capital taxes due to the pressure of tax competition. We therefore obtain that the median voter favors a partial degree of decentralization. The equilibrium degree of decentralization is increasing in inequality and the transferability of public funds, and decreasing in capital productivity. We therefore obtain a theory of decentralization despite the absence of spillovers or heterogeneity in tastes across districts.

*We are grateful to Ernesto Dal Bó, Ken Shotts, Romain Wacziarg, John Wallis, and Barry Weingast for helpful discussions. This is a preliminary working draft; please do not cite or distribute without permission. Any comments or suggestions are welcome and may be emailed to hatfield@stanford.edu or g.padro@lse.ac.uk.
1 Introduction

A basic feature of the institutional framework of a country is what level of government responsible for the provision of each publicly provided good. As it has been recognized in the fiscal federalism literature, incentives and constraints of policy-makers vary substantially with different allocations of such decision rights.\(^1\) In most countries, provision rights are partially decentralized, with some goods provided and funded at the local level while other goods are provided at the central level. In the US, for instance, education is mostly funded and the local level, using real state taxes, while spending in parks or highways is mostly decided at the federal level. Spillovers across districts, local differences in tastes and economies of scale have been used in the fiscal federalism literature to conform theories of what goods should be provided at which level.

Tax competition is a recurrent concern in this literature. Oates (1973) first articulated the possibility that fiscally federal countries may suffer from pernicious tax competition between subnational units for the location of mobile factors of production such as capital. According to this intuition, the decentralization of the public provision of consumptive goods reduces welfare as local units engage in a “race to the bottom,” reducing taxes on mobile factors – and accordingly reducing provision levels – in order to attract capital from other districts.\(^2\) It follows from most of the models in this literature that a central authority would do better because it does not face as strong a competitive pressure and would therefore be able to use the optimal mix of taxes. These results depend on three assumptions. First, governments maximize aggregate surplus within their jurisdictions. Second, governments can commit to a fixed level of taxation. Finally, the total capital stock is fixed, albeit mobile across districts.

We revisit the tax competition results using a framework that relaxes these three assumptions. As a result, we obtain a rationale for the partial decentralization of public good provision that is completely different from the arguments put forth in the literature. In particular, we argue that decentralization serves as a way of committing to a low level of capital taxation and we show that such commitment might be desirable even for agents in society that hold little capital. We therefore find that the effects of tax competition can be beneficial.

---

\(^1\) For comprehensive literature review on federalism, see Weingast (2006)
Our argument is as follows. If there is inequality in capital ownership among citizens, the political process can lead to an excessive use of capital taxes for two reasons. First, holding public provision constant, the median voter might prefer to use capital taxes even in the presence of non-distortive lump-sum taxes. In particular, if the median voter holds less capital than the average citizen, the use of capital taxes allows her to avoid part of her proportional tax burden. Second, given this unequal distribution of tax burden, the median voter will also prefer an excessive level of public good provision.\footnote{Besley and Coate (1991) already show that public provision of goods can serve as a redistribution device.}

If the capital stock is fixed as assumed in the literature, this political economy reason for excessive capital taxation only results in a distributional change, with no aggregate efficiency effect. This is, however, not realistic. The amount of capital available in the economy at a given time depends on previously taken savings and investment decisions. In this case, if agents expect a high level of capital taxation, the aggregate amount of capital in the economy is adversely affected. This contraction not only constitutes an aggregate inefficiency but it also hurts the median voter as there is less surplus to be redistributed. It follows that the median voter would like to commit to a lower level of capital taxes before agents take their capital generation decisions.

Since budgetary decisions are taken every year, voters lack the capacity to commit to a given level of capital taxation. In contrast, the institutional structure of the nation is much more difficult to change and hence it provides some commitment technology. In particular, decentralization serves to reduce the expected aggregate capital taxes. The reason is precisely tax competition: the need to attract factors of production forces local governments to avoid high capital taxation. By decentralizing the provision and funding of some goods the median voter faces gains and losses. On the one hand she can commit to a lower level of capital taxation. On the other hand she loses the ability of using the provision of decentralized goods as redistributive tools in her favor. As a consequence, the median voter favors a partial degree of decentralization: despite the fact that public goods are ex ante identical, she wants a fraction of them provided at the local level.

We illustrate this logic with a parsimonious model. In our model there is a single factor of production, capital, which is mobile and costly generated. Agents are identical in tastes, but they differ in their capacity to generate capital, and hence capital holdings are unequal. We endow the different administrations (a central gov-
ernment and a discrete number of identical districts) with two tax instruments: a tax on capital holdings and a lump-sum tax. There is a continuum of public goods, a fraction of which are to be funded and provided at the local level. Voters then decide on the level and type of taxation and public provision. In equilibrium, we show that the goods provided by the central government are funded using capital taxes and are oversupplied, while goods provided by the districts are funded using head-taxes and therefore are supplied at the optimal level due to the pressure of tax competition. As a consequence, increasing the degree of decentralization automatically decreases the equilibrium level of capital taxation.

We then proceed to examine a version of the model that allows us to solve for the degree of decentralization preferred by every voter in the economy. We show that the optimal degree of decentralization from the point of view of the median voter is increasing in the transferability of public funds through public good provision, almost always increasing in inequality and decreasing in the productivity of capital.

When we allow public goods to be heterogeneous in transferability, we obtain a striking result. All voters agree on the ordering in which public goods should be decentralized. Therefore, set of Pareto optimal decentralization schemes easy to characterize. In particular, we show that public goods that allow high transferability of public funds, i.e. those with slowly decreasing marginal returns to investment, are the best to decentralize as they are the most oversupplied if provided by the central government. The only point of disagreement between voters is over the amount of power that is to be devolved to the states; the richer the agent is, the more decentralization he desires. It thus follows from our analysis that political parties that represent capital owners should favor increased levels of decentralization, as the Republican Party does in the US.

This characterization of decentralization provides a new lens with which to consider the issue of what goods to be centralized. Both the seminal work of Oates (1972) and the more recent work by Besley and Coate (2003) and Lockwood (2002) find that the optimal decentralization structure depends on the level of taste disparity and the degree of spillovers between districts. Our work abstracts from both these concerns as it assumes taste homogeneity across districts and no spillovers. However, we still find reasons for partial decentralization: to change the tax base used in equilibrium, and in so doing, change the incentives for capital investment and the degree of redistribution generated by public spending. In addition, to our knowledge our framework is the first one that allows for a parsimonious analysis of the degree of decentralization.
Our insight is closely related to the classic dynamic inconsistency problem in capital taxation in the macroeconomics literature. Indeed, in the context of monetary policy, Rogoff (1985) and Kehoe (1989) show that policy coordination between countries might not be desirable. The reason is akin to the main force behind the beneficial effects of tax competition between jurisdictions in our model: by forcing policy-makers to compete, excessive taxation levels cease to be equilibrium outcomes. We contribute to this literature by showing that partial federalism provides a natural way of trading off the good and the bad effects of competition.

The key point of our argument is that institutions are more resilient to change than policies. While Congress can not bind the hands of future Congresses, the construction of a constitutional structure is binding, and changes in constitutional structure are very difficult. Federalism, in particular, is an institutional feature that can easily be self-sustaining. A Federal structure thus provides a commitment technology much stronger than the mere promise not to increase capital taxes. In the language of mechanism design, federalism allows commitment to the allocation of decision rights, but not to decisions themselves. Our work, then, is part of an emerging literature on second-generation federalism, as is discussed in Oates (2005) and Inman and Rubinfeld (1997). This literature models political institutions explicitly, and does not rely on politicians to act as social planners. This approach leads to new insights into how different governmental forms will lead to different economic outcomes.

The next section of this paper describes the general model of partial decentralization that we propose. Section 3 provides the analysis of the model and discusses the main intuitions of the paper. The following section takes some particular functional form assumptions in order to examine the degree of decentralization favored by each voter and provide rich comparative statics. Section 5 provides an extension of the model to heterogeneous public goods. The last section concludes. All proofs are in the Appendix.

---


5 For an argument why federalism can be self-sustaining in equilibrium, see de Figueiredo and Weingast (2005) and Weingast (2006) and references contained therein.

6 In the context of our model, a Constitutional clause fixing a maximum rate of capital taxation would also work. Such clauses are, however, virtually absent in practice.
2 The General Model

We consider a static economy where agents first decide how much to invest in capital, then governments decide upon tax and spending policy, then agents decide where to invest their capital, and finally production and consumption are realized. We start by describing the physical and institutional environment, followed by the economic opportunities of agents and their preferences.

Assume that the economy is divided into \( J \) identical districts, each with its own local government, and each with a total mass of individuals \( 1 \). There are two levels of government, the local level (which can capture district, municipality or state level) and the central government (or federal level). These administrations are assumed to use their revenues to provide publicly provided goods to the citizens within their jurisdiction. There is a continuum of size \( 1 \) of such homogeneous publicly provided goods. District governments are responsible for providing goods \([0, \lambda]\), and the national government is responsible for providing on goods \((\lambda, 1]\). Since the higher \( \lambda \) is, the more responsibilities districts take on, \( \lambda \) constitutes a measure of the degree of decentralization in this economy.

Administrations must raise revenues to meet their expenses in publicly provided goods. We consider a simple economy with a mobile factor of production, \( k \) and a non-mobile factor which we can associate with land. Hence, governments have access to both a tax on capital within their jurisdiction and a land tax which amounts to a lump-sum tax.\(^7\) Denote by \( \tau \) and \( T \) the tax level on capital and land levied by the central government and by \( \tau_j \) and \( T_j \) the level of taxation levied by district \( j \).\(^8\) We shall assume that land taxes are constrained to be nonnegative. Denote by \( s(p) \) the amount of spending on publicly provided good \( p \in [0, 1] \) by the administration responsible for its provision (we denote by \( s_j(p) \) the level of spending on good \( p \) by district \( j \)). With this notation, the budget constraint for the district government is given by

\[
\int_0^\lambda s_j(p) \, dp = \tau_j k_j + T_j \tag{1}
\]

where \( k_j \) denotes the amount of capital invested in district \( j \). The budget constraint

\(^7\)To ensure the existence of a Condorcet winner in the policy space, we constrain the tax on capital to be linear. See Romer (1975), Roberts (1977), and Meltzer and Richard (1981).

\(^8\)Note that our results do not change if \( \tau_j \) and \( \tau \) are taxes on capital returns and not capital investment.
for the national government is given by

$$\int_{\lambda}^{1} s(p) \, dp = \tau k + T$$

(2)

Consumption goods are produced by a continuum of firms at the local level using capital and land. Denote by $F(k_j)$ the production function. $F(k_j)$ is increasing, weakly concave, and smooth. Since it does not necessarily display constant returns to scale, we assume that land accrues the returns that are not captured by capital owners. Such returns to land are shared equally by residents of district $j$.\(^9\) Denote by $\rho_j$ the pre-tax rate of return to capital in district $j$. Competition within districts implies that capital captures its marginal contribution to production, or

$$\rho_j = F'(k_j)$$

(3)

for each district. Moreover, since capital is assumed to be perfectly mobile across districts, its after-tax returns must be equalized. It follows that

$$r = \rho_j - \tau_j - \tau$$

(4)

where $r$ is the net return to capital and in equilibrium it is uniform throughout the economy.

We can now proceed to describe the preferences of agents and their economic opportunities. An agent $n$ in district $j$ is characterized by a type, $\beta^n$, which determines how costly investment in capital is for the agent. This type captures inequality given by different asset levels which might provide different access to credit, or simply skill, social contacts or educational opportunities. Types are distributed according to some cumulative distribution function $H(\cdot)$ on $[\bar{\beta}, \tilde{\beta}]$. Agent $n$ needs to pay a cost $v(k^n; \beta^n)$ to generate an amount $k^n$ of productive capital. $\nu(\cdot)$ is a smooth function that is increasing and convex in its first argument and $-\nu(k^n, \beta^n)$ satisfies the single crossing property. Agent $n$’s preferences are thus given by

$$u(c^n, s(j), k^n; \beta^n) = c^n_j + \int_{0}^{1} G(s(p)) \, dp - v(k^n; \beta^n)$$

where $c^n$ denotes agent $n$’s consumption and $G(\cdot)$ is a smooth, increasing, and concave

\(^9\)This assumption is made for simplicity and it amounts to assuming that all residents own an equal share of land in the district. This assumption allows us to focus on different capital holdings as the sole source of inequality in this economy.
function of spending in the publicly provided goods that the agent will enjoy. For an agent living in district \( j \), \( s(p) = s_j(p) \) for goods \([0, \lambda]\), while goods \((\lambda, 1]\), being provided by the central government, receive \( s(p) \) that is equal across districts. Given the taxation patterns described above, agent \( n \) in district \( j \) enjoys consumption equal to

\[
e^*_n = r k^*_n + F(k_j) - \rho_j k_j - T_j - T
\]

where \( r k^*_n \) are the net returns to her capital holdings, which she can invest anywhere in the economy, while \( F(k_j) - \rho_j k_j - T_j - T \) are the net returns to her land holdings which are limited to district \( j \).

The timing of the model is as follows.

1. Each agent \( n \) in each district \( j \) decides how much productive capital to generate, \( k^*_n \).

2. By simple majority, taking the rate of return on capital \( r \), and their budget constraint as given, the citizens in each district choose the Condorcet winner in their policy space \((\tau_j; T_j; s_j(p), p \in [0, \lambda])\).

3. By simple majority, taking the budget constraint as given, all the agents choose the Condorcet winner in the policy space of the national government \((\tau; T; s(p), p \in (\lambda, 1])\).

4. After observing taxation patterns across the economy, agents decide in which district to invest their productive capital, \( k^*_n \).

This model has a number of noteworthy features. First, note that the publicly provided goods we discuss here do not strictly correspond to public goods. In the utility for publicly provided goods, \( G(\cdot) \), only the spending associated with the district government features for goods \([0, \lambda]\) and hence a citizen of district \( j \) does not obtain any utility from resources spent by district \( k \) in these goods. With this, we abstract from cross-district spillovers and the public goods problem, and we focus on locally provided goods such as education or health.\(^{10}\)

Second, we use a quasilinear utility function for consumption and publicly provided goods. With this assumption, we ensure that there are no income effects in the enjoyment of such goods. Therefore, the tension to decide on the level of public spending is exclusively given by the fact that more productive agents will bear a higher level of taxation and therefore will favor a lower level of spending. In this

\[^{10}\text{Goods} (\lambda, 1], \text{provided by the central government, could actually be pure public goods. We are agnostic about this as budget constraint (2) works with either type of goods.}\]
sense, public provision of goods serves as a redistributive tool because its enjoyment is equalized within district while its funding, insofar capital taxes are used, depends on the level of productive capital generated by each individual.

2.1 Definition of Equilibrium

For a given level of decentralization \( \lambda \), a subgame perfect equilibrium is a capital investment decision function \( k (r, \beta) \) for \( \beta \in [\bar{\beta}, \bar{\beta}] \), policy decisions \( T, \tau, s(p) \) for \( p \in [\lambda, 1] \), \{\( T_j, \tau_j, \rho_j \)\} \( j = 1, ..., J \), \{\( s_j(p) \) for \( p \in [0, \lambda] \)\} \( j = 1, ..., J \), an after-tax rate of return on capital \( r \), and investment location decisions such that

1. Capital markets are perfectly competitive both intra- and interdistrict: \( \rho_j = F'(k_j) \) in each district and \( r \), the after tax rate of return in each district, is \( \rho_j - \tau - \tau_j \).

2. The districts, taking the rate of return on capital \( r \) and their budget constraint as given, choose the Condorcet winner in their policy space.\(^{11}\)

3. The national government, taking its budget constraint as given, chooses the Condorcet winner in its policy space.\(^{12}\)

4. Agents choose to generate an amount of capital \( k (r, \beta), \beta \in [\bar{\beta}, \bar{\beta}] \) to maximize their utility.

We note that, as in many models where the tax base reacts to expected taxation levels, there may be multiple equilibria. In particular, there may exist equilibria on the wrong side of the Laffer curve, where the total revenue of the government is locally decreasing in the level of taxation. In such situations, we only consider equilibria in the increasing side of the Laffer curve, i.e. that a small increase in the (expected) capital tax rate yields an increase in government revenues. We call such equilibria, standard equilibria.

\(^{11}\)It will be shown that such a Condorcet winner exists.

\(^{12}\)See previous footnote.
3 Characterization of Equilibrium

3.1 The Problem of the Agent

At stage 1, the agent must decide how much capital to generate. The problem of agent \( n \) in district \( j \) is to solve

\[
\max_{k^n_j} \{ u(c^n, s(j), k^n, \beta^n) \}
\]

Thanks to the quasilinear structure of preferences, this problem reduces to

\[
\max_{k^n_j} \{ rk^n_j - \nu(k^n_j, \beta^n) \}
\]

for an equilibrium-consistent expected return \( r \). Due to the single-crossing property of the objective function, it is immediate that higher types will generate more capital. It is also immediate from the agent’s maximization problem that the agent will, regardless of type, generate more capital the higher the expected after-tax rate of return.

Let \( k(\beta^n, r) \) be the amount of capital generated by an agent of type \( \beta^n \) who expects an after-tax rate of return on capital \( r \). We denote by

\[
k(\beta, r) \equiv \int_{\beta}^{\beta^n} k(\beta, r) \ dH(\beta)
\]

the amount of capital generated in a district when the expected net return is \( r \). Also, we let \( k_{\text{med}} \equiv k(\beta_{\text{med}}, r) \) be the amount of capital held by the citizen with the median capacity to generate capital, \( \beta_{\text{med}} \). We shall assume that

\[
1 < \Phi(r) \equiv \frac{k(r)}{k_{\text{med}}} < \infty
\]

for all \( r \). This is essentially the standard assumption that the median voter (it will be shown that it is the voter with the median \( \beta \)) has less capital than the average agent.\(^{13}\)

From the equilibrium condition on capital markets (4), we have that the after-tax return on capital is simply the marginal return to capital minus the national and local capital taxes. Using this condition, the following result can be shown:

\(^{13}\)See, for instance, Persson and Tabellini (2000) and references contained therein.
Lemma 1 The amount of capital generated by agent \( n \) is decreasing in the level of capital taxes.

The intuition behind this result is standard: as capital taxes increase, the after-tax return to capital falls and therefore agents generate less capital in response. This is at the core of a time inconsistency problem for taxation agencies. Governments would like to promise to keep capital taxes low \textit{ex ante} in order to stimulate capital generation. However, after capital has been generated, it is perfectly efficient to tax capital to provide for public goods since capital is not immediately destroyed \textit{ex post}.

### 3.2 The Problem of the District Government

To solve the decision problem of the district government, we shall first find the tax policy favored by a given agent \( n \) with type \( \beta^n \). Afterwards, we shall show that a Condorcet winner for the district as a whole exists. Recall that the district government decides on policy after capital has been generated, but before agents have chosen in which district to invest that capital. Hence, the problem of voter \( n \) in district \( j \) is:

\[
\max_{k_j, T_j, \tau_j, s_j(p)} \left\{ c^n_j + \int^\lambda_0 G(s_j(p)) \, dp + \int^1_0 G(s(p)) \, dp - \nu(k^n_j, \beta^n) \right\}
\]

which, using (5), simplifies to

\[
\max_{k_j, T_j, \tau_j, s_j(p)} \left\{ F(k_j) - \rho_j k_j - T_j + \int^\lambda_0 G(s(p)) \, dp \right\}
\]

subject to the district budget constraint

\[
\int^\lambda_0 s_j(p) \, dp = \tau_j k_j + T_j
\]

and the constraint from capital mobility

\[
r = \rho_j - \tau_j - \tau
\]

Looking at program (6) it becomes obvious that within the district, there is no political conflict. Since capital is mobile and can escape high levels of taxation in a given district, residents of that district only care about maximizing returns to land, \( F(k_j) - \rho_j k_j \) and obtaining publicly provided goods. Therefore, \( \beta^n \) drops out of the problem, and all agents have the same preferences at the district level: the Condorcet
The winner will be the policy most preferred by every single agent.

Note also that it is immediate from the concavity of $G$ that $s_j(p) = \frac{\tau_j k_j + T_j}{\lambda}$.

Plugging this condition in the objective function we obtain
\[
\max_{k_j, T_j, \tau_j, s_j(p)} \left\{ F(k_j) - \rho_j k_j - T_j + \lambda G \left( \frac{\tau_j k_j + T_j}{\lambda} \right) \right\}
\]

Now, taking the first order condition with respect to $T_j$ it is immediate that
\[
G' \left( \frac{\tau_j k_j + T_j}{\lambda} \right) = 1
\]
which implies the Samuleson condition: each good is provided at the efficient level given that due to the availability of a lump-sum tax the opportunity cost of public funds equals 1. Now, taking the first order condition with respect to $k_j$ and plugging in (7) one obtains
\[
F'(k_j) - \rho_j + \tau_j = 0
\]
which, using perfect competition for capital within districts (3), immediately implies that $\tau_j = 0$. Therefore publicly provided goods at the district level are entirely funded by lump-sum taxes. Essentially, the district is trying to maximize profits at the local level buying capital from a competitive market for capital at price $r + \tau$. Efficiency then requires that capital is rented up to the level where $F'(k_j) = r + \tau$. If the district taxes capital at any positive level, capital mobility condition (4) requires that the district obtains less capital than it would be efficient. Therefore, it is optimal to refrain from capital taxation and raise revenues using exclusively lump-sum taxes.

We have established the following proposition:

**Proposition 1** For any level of decentralization $\lambda$, each district government will efficiently invest in each public good (i.e. $G'(s_j(p)) = 1$ for all $j = 1, ..., J, p \in [0, \lambda]$) using only a land tax.

### 3.3 The Problem of the Central Government

Because capital cannot escape taxation at the central level, federal capital taxes change net returns to capital. Tax incidence is therefore unequal across voters if capital taxes are used and hence there is potential for political conflict. We shall first consider the problem of optimal policy from the perspective of the median agent – i.e., the agent with type $\beta^{med}$ – and then later show that the median agent’s ideal point is indeed a Condorcet winner within the policy space of the central government.
The problem for the median agent in some district \( j \) is

\[
\max_{T, \tau, s(p)} \left\{ c_j^{\text{med}} + \int_0^\lambda G(s_j(p)) \, dp + \int_\lambda^1 G(s(p)) \, dp - \nu \left( k_j^{\text{med}}, \beta^{\text{med}} \right) \right\}
\]

which, given that capital generation decisions have already been taken, simplifies to

\[
\max_{T, \tau, s(p)} \left\{ -\tau k^{\text{med}} - T + \int_\lambda^1 G(s(p)) \, dp \right\}
\]

subject to the national government’s budget constraint

\[
\int_\lambda^1 s(p) \, dp = \tau k + T
\]

Again, it is immediate from the concavity of \( G \) that \( s(p) = \frac{\tau k + T}{1-\lambda} \).

It is also important to note that the central government takes the total amount of capital \( k \) as fixed. This is in contrast with the district governments that face capital flight should they decide to tax capital. It follows that, for the central government, both the head tax and the capital tax are completely nondistortive \( \text{ex post} \). However, the median agent is not indifferent. If lump-sum taxes are used, the agent with \( \beta^{\text{med}} \) bears an equal burden of taxation with any other agent in the economy. Conversely, since \( \Phi \equiv \frac{k}{k^{\text{med}}} > 1 \), when capital taxes are used the median agent agent pays less taxes than the average agent. Therefore, from the perspective of the median agent, it is better to use the capital tax for any level of public provision of goods.

Furthermore, precisely because the agent with \( \beta^{\text{med}} \) pays a smaller share of taxes, her ideal level of public provision of goods is higher than what efficiency requires. In particular, we find that the solution to the median agent’s ideal policy is

\[
G' \left( \frac{\tilde{\tau} k}{1-\lambda} \right) = \Phi^{-1} < 1
\]

\[
\tilde{T} = 0
\]

Recall that the efficient level of public provision is set at \( G' = 1 \) due to the fact that preferences are quasilinear. Hence, condition (8) implies that an excessive level of public provision of goods is preferred by the median agent. Essentially, the median agent wants a higher level of public spending because she is using public provision of goods as a channel of redistribution from large owners of capital to small owners. Note that voters with less capital than the median agent agree with the policy to not use a head tax. The only difference is they would prefer even more public spending, as
redistribution via public provision is even more appealing to them. Moreover, if only a capital tax is available, then voters with more capital than the median agent will prefer a lower capital tax and a less than efficient level of public provision.

To see that the policy preferred by the median agent is indeed a Condorcet winner, consider figure 1. Given a set amount of revenue, it is clear that all agents will wish to spend it in the most efficacious way on public goods. Given this, the real question is over the level of \( T \) and \( \tau \). Consider any policy other than that most favored by the median voter. If it involves a higher capital tax, as well as a positive head tax (a policy in region A of figure 1), then the policy with the same capital tax and no head tax will be favored by all the voters; at its turn, this policy is dominated by the policy with no head tax and a capital tax \( \hat{\tau} \) in the eyes of all agents with an amount of capital greater than or equal to the median amount of capital. Hence, by individual transitivity of preferences, \((\hat{\tau}, 0)\) is preferred to any alternative policy in region A. If the alternative policy has a higher total level of funding, but a lower capital tax (a policy in region B of figure 1) then the policy that holds funding constant, while decreasing the head tax and increasing the capital tax to \( \hat{\tau} \) is favored by all voters with less capital than the mean. At this point, reducing the head tax to 0 is favored by all voters, since with a capital tax of \( \hat{\tau} \) each public good has a marginal return of less than one. Hence, again by individual transitivity of preferences, \((\hat{\tau}, 0)\) is preferred to any policy in region B. Finally, consider a policy with a lower total level of funding than \((\hat{\tau}, 0)\) (a policy in region C). Again, decreasing the head tax to zero and increasing the capital tax while holding funding constant is favored by all voters with less capital than the mean. After that, increasing the capital tax is favored by all voters with less than or equal to the capital of the median voter. Hence, by individual transitivity of preferences, \((\hat{\tau}, \hat{T} = 0)\) is preferred to any policy in region C.

We summarize this result in the following proposition:

**Proposition 2** The national government will use only capital taxes, and will set \( G'(\frac{r_k}{1-\lambda}) = \Phi^{-1} \), which provides more than the efficient amount of the public good.

### 3.4 Optimal Capital Taxation

We have established that ex post, only capital taxes are used to fund public provision of goods by the central administration. However, the implied level of capital taxes might be too high from an ex ante perspective, even from the point of view of the median agent. Ex post, the median voter only considers the tradeoff between her
individual lost consumption and public provision when deciding how much to tax in order to finance the public good, and does not consider the effect on both his own and others’ incentives to invest \( \text{ex ante} \). In other words, due to her inability to commit to low capital taxes, the median agent finds herself with a smaller capital pool from which to redistribute. If she could commit, she would solve the following program, taking district decisions as given

\[
\max_{\tau} \left\{ (F'(k) - \tau) k_{\text{med}} + (F(k) - F'(k) k) - T_j \right\}
\]

\[
+ \lambda G \left( \frac{T_j}{\lambda} \right) + (1 - \lambda) G \left( \frac{\tau^k}{1 - \lambda} \right)
\]

\[-v \left( k_{\text{med}} \right)
\]

Note that we are constraining this program already to \( \hat{T} = 0 \). The first term in this program, \((F'(k) - \tau) k_{\text{med}}\) corresponds to the returns to capital that the median voter will perceive, while the second term \((F(k) - F'(k) k) - T_j\) captures the returns to her landholdings. The first order condition yields

\[
kG' \left( \cdot \right) - k_{\text{med}} + \left( \tau G' \left( \cdot \right) - F'' \left( k \right) \left( k - k_{\text{med}} \right) \right) \frac{\partial k}{\partial \tau} = 0
\]

The first term captures the marginal benefit of an additional unit of public provision, while the second term relates to the capital tax burden on the median agent. We have, however and additional term that must be strictly negative, as \( G \left( \cdot \right) \) is increasing, \( F \left( \cdot \right) \) is concave, and we know from Proposition 1 that \( \frac{\partial k}{\partial \tau} \) is smaller than 0. The first term within the parentheses, \( \tau G' \left( \cdot \right) \) is the additional utility the median agent receives from the fact that an increase in capital holdings increases the tax base and hence allows for more spending in public provision. The second effect is the private consumption effect: as more capital is generated, average returns to land increase, while private returns to capital decrease at the margin. Since the median voter holds less capital than average, she favors such increase. (Note that the envelope theorem ensures that the effect of changing \( \tau \) on the median voter’s utility through the change in \( k_{\text{med}} \) is zero.)

By comparing with expression from the equivalent first order condition from the preceding subsection, we have that only the first to terms were present.

\[
kG' \left( \frac{\tau k}{1 - \lambda} \right) - k_{\text{med}} = 0
\]

This implies that the solution to the preceding subsection induces a higher level of capital taxation that is optimal when the median agent is able to commit to a level
of capital taxation \textit{ex ante}. Hence we have the following result:

\textbf{Lemma 2} \textit{Imposing }$\mathcal{Y} = 0$, \textit{the median voter’s preferred capital tax rate when the national government can precommit is lower than the Condorcet winning capital tax rate when the government can not precommit.}

The median agent, then, would like to commit to lower taxes in order to promote capital generation.

\section*{3.5 Capital Taxation and Decentralization}

In section 3.2 we have shown that district governments do not use capital taxes to fund the provision of goods that fall within their jurisdiction. The following proposition establishes that by increasing the amount of goods that district governments are responsible for, the expected level of capital taxes must fall.

\textbf{Proposition 3} \textit{In any standard equilibrium, the equilibrium capital tax rate }$\tau + \tau_j$\textit{ is a decreasing function of decentralization:}

$$\frac{d(\tau + \tau_j)}{d\lambda} = \frac{d\tau}{d\lambda} < 0$$

This follows straightforwardly from our earlier results on capital taxation.

The previous subsection establishes that the median voter would desire commit to lower capital taxes. Proposition 3 shows that decentralization provides an institutional way to commit to such policy. In most countries, the distribution of responsibilities between the different layers of government are either enshrined in the constitution or require a major legislative effort to change. In contrast, taxation policy is typically decided every year, during the budgetary process and hence is subject to the commitment problem we have highlighted here. Institutional arrangements can not be so finely tailored as budgetary policy, but this lack of flexibility comes with an important benefit: the electorate is able to precommit to institutional arrangements in a way they can not with respect to policy. Insofar the allocation of public provision of goods between local and central governments is an institutional feature, it provides as a consequence a strong commitment device to a reduced level of capital taxation.

By the argument above, the median agent is willing to commit to a high level of decentralization \textit{ex ante}. However, typically complete decentralization will not be optimal from her perspective: every good that is transferred from the central government to the district governments cannot be used as a redistributive device from
high capital owners to small capital owners. Also, the level of provision of such good is reduced. Therefore, the median voter typically prefers an interior solution to the constitutional problem that entails a partial level of decentralization.

In the following section we assume specific functional forms that allow us to explicitly solve for the level of decentralization preferred by the median voter. In such an example we can explore how the preferred level of decentralization varies with factors such as productivity, inequality, or the redistributive nature of publicly provided goods.

4 The Equilibrium Level of Decentralization

We consider a particular case of the model developed above. Assume that technology is linear in $k$, $F(k) = Ak$, where $A$ captures the general level of productivity in this economy. Moreover, let $v(k^n; \beta^m) = \frac{(k^n)^2}{2\beta^m}$. Finally, assume that $G(s(p)) = \left[\frac{s(p)^\alpha - 1}{\alpha}\right]$, for $\alpha < 1$. We consider a distribution of types $H(\cdot)$ such that the expected value, $\tilde{\beta}$ is greater than the median value, $\beta^{med}$. It is straightforward to see that these functional assumptions satisfy the conditions of the general model. We do now proceed to solve this case.

The problem of agent $n$ in district $j$ is particularly simple.

$$\max_{k^n} k^n(A - \tau - \tau_j) - T - T_j - \frac{(k^n)^2}{2\beta^n}$$

and yields a linear solution,

$$k^n = \beta^n(A - \tau - \tau_j) \tag{9}$$

The government of the district also faces a simple problem given that with linear technology there are no returns to land.

$$\max_{T_j, \tau_j} - T_j + \lambda \left(\frac{T_j + \tau_j}{\lambda}\right)^\alpha - 1$$

This objective function already assumes that public spending will be equally distributed across the $\lambda$ goods districts are responsible for. Note also that with linear technology, the capital mobility constraint is particularly tight. In particular, district $d$ only receives any capital investment if $\tau_d = \min_{j \in J} \tau_j$. Not surprisingly, Bertrand competition between districts yields $\tau_j = 0 \forall j \in J$. Hence, the solution to the problem of
the district is \((\tau_j, T_j) = (0, \lambda)\).

The problem of the median voter for central government policies is slightly more involved. Integrating (9) we obtain the average level of capital as a function of taxation. In particular, 

\[
\bar{k} = \int \beta (A - \tau C) h(\beta) d\beta = (A - \tau C) \bar{\beta},
\]

where \(\bar{\beta}\) is the expected value of the distribution \(H(\cdot)\) of abilities to generate capital. With this, we can write

\[
\max_{T, \tau} k^{\text{med}} (A - \tau) - T + (1 - \lambda) \left( \frac{T + \tau \bar{k}}{1 - \lambda} \right)^\alpha - 1
\]

subject to \(T \geq 0\). A little algebra shows that this constraint must be binding as long as \(\frac{\bar{k}}{k^{\text{med}}} > 1\). With linear technology, we obtain

\[
\frac{\bar{k}}{k^{\text{med}}} = \frac{\bar{\beta}}{\beta^{\text{med}}} \equiv \Phi > 1
\]

which immediately implies that no lump-sum taxes are used. Taking the first order condition with respect to \(\tau\) yields

\[
\tau = \Phi \gamma \frac{1 - \lambda}{k^{\text{med}}}
\]

where \(\gamma = \frac{\alpha}{1 - \alpha}\).

The equilibrium level of capital taxes and capital generation can be found by solving the non-linear system of equations (9) and (10). By doing so, one obtains a well-defined capital generation function \(k^{\text{med}}(\beta^{\text{med}}, \alpha, \Phi, \lambda)\) for the standard equilibrium of this model.

**Proposition 4** The linear-quadratic model of section 4 admits a unique standard equilibrium. This equilibrium features the following comparative statics:

- \(\frac{\partial k^{\text{med}}}{\partial \lambda} > 0\)
- \(\frac{\partial k^{\text{med}}}{\partial \alpha} < 0\)
- \(\frac{\partial k^{\text{med}}}{\partial \Phi} < 0\)
- \(\frac{\partial k^{\text{med}}}{\partial A} > 0\)

Since \(\tau\) is common for all agents, the equilibrium capital level of any agent \(n\) is proportional to \(k^{\text{med}}\). In particular, \(\frac{k^n}{k^{\text{med}}} = \frac{\beta^n}{\beta^{\text{med}}}\). Hence the comparative statics in Proposition 4 are common for all agents, and also for the aggregate level of capital.
generated in the economy. As was found in the general model, an increase in decentralization $\lambda$, reduces the level of capital taxation expected by the agents therefore fostering the generation of production capital. Holding $\lambda$ constant, however, we find interesting comparative statics. In particular, as $\alpha$ increases, capital reduces. The intuition behind this result is closely related to the use of publicly provided goods as redistribution devices. As $\alpha$ increases, the returns to such goods become closer to linear and therefore diminishing returns take longer to set in. As a consequence, the median voter, at the time of deciding the amount of provision desired, will be tempted to increase taxation—and consequentialy redistribution—. These expectations depress capital generation. Similar intuition lies behind the result that capital is decreasing in $ex\ ante$ inequality, $\Phi$. As inequality increases, the temptation to expand public spending in order to redistribute becomes stronger. In such circumstances, expected capital taxation increases thereby reducing incentives to generate capital. Finally, an increase in productivity increases returns to capital. Since in this linear model $ex\ post$ inequality is equal to $ex\ ante$ inequality, $\Phi$, these excess returns are not taxed away and therefore capital reacts positively.\(^{14}\)

### 4.1 The Constitutional Problem

Now consider the following timing:

1. By simple majority, a Constitution is voted such that the level of decentralization $\lambda^*$ is fixed.

2. All agents in all districts generate their productive capital

3. By simple majority, districts and central government choose the Condorcet Winners in their respective policy spaces

4. Agents allocate their productive capital across districts

5. Production and consumption occurs

Therefore in this section we explicitly consider how the median agent balances her desire of redistribution with her commitment problem. Consider a subgame perfect

\(^{14}\)Note that the fact that inequality is fixed at $\Phi$ is a convenient property of the functional forms assumed in this section. In the general model, it might easily be that $\frac{k}{k_{med}}$ differs from $\frac{2}{\sqrt{\Phi}}$. In such a case, the capital ratio typically moves in ways that complicate some of the comparative statics. For instance, if $\frac{k}{k_{med}}$ increases as aggregate capital expands, capital taxation also increases leading to an adverse effect. If productive agents fear that all their capital retuns will be taxed away, it is possible that capital does not react at all to seemingly positive shocks to the economy.
equilibrium in which at stage 1 the median agent takes as given the investment and taxation decisions that will occur afterwards. In that case, her program can be written as

$$\max_k k^{med}(A - \tau) - T_j + (1 - \lambda) \left( \frac{r_k}{1 - \lambda} \right)^{\alpha} - 1 + \lambda \left( \frac{T_k}{\lambda} \right)^{\alpha} - 1 - \frac{(k^{med})^2}{2\beta^{med}}$$

By substitution of the results in the previous subsection, we can further simplify this expression to

$$\max_k A k^{med} - \frac{(k^{med})^2}{2\beta^{med}} - \Phi^\gamma (1 - \lambda) - \lambda + (1 - \lambda) \frac{\Phi^\gamma - 1}{\alpha}$$

where the first two terms correspond to her production incentives (gross returns to capital minus costs) and the next terms correspond to taxation and redistribution. Note in particular that as $\lambda$ increases, the last two terms hurt the median agent as lump-sum taxes go up and public provision of goods diminishes. The first three terms capture, however, the positive effects of such increase in $\lambda$ as her own generation of capital will react positively due to the reduction of capital taxes. We can now state the main result on optimal decentralization.

**Proposition 5** In the model of section 4 there exists a unique optimal level of decentralization for the median agent, $\lambda^*$ in a standard subgame perfect equilibrium. Such optimal level of decentralization features the following comparative statics

- $\frac{\partial \lambda^*}{\partial \alpha} > 0$
- $\frac{\partial \lambda^*}{\partial A} < 0$

Uniqueness comes from the fact that in a standard equilibrium, $k^{med}$ is a concave function of $\lambda$. These comparative statics are in line with the intuitions generated in Proposition 4. As the degree of transferability in public provision increases, the temptation to use public provision as a redistribution mechanism is exacerbated. As a consequence, the equilibrium level of capital is greatly reduced due to the expectation of high taxes. To avoid such reaction, the median agent prefers to vote for a constitutional arrangement that allocates more responsibilities to the local levels of government, thereby ensuring capital owners that taxes will be lower. Conversely, an increase in productivity $A$ increases the incentives to generate capital. In that case, the median agent can afford to reduce the level of decentralization in order to tax and redistribute some of the returns to the extra capital generated.
The comparative statics with respect to inequality $\Phi$ are, however ambiguous. In most of the cases, $\frac{\partial \lambda}{\partial \Phi} > 0$ as greater inequality also increases the temptation to tax capital. In this natural case, the intuition developed above applies. However, for small $\alpha$ and $\Phi$, the sensitivity of $\tau$ to an increase in inequality is quite small, as expression (10) shows. Since the reaction of aggregate capital to a change in $\Phi$ is muted in these circumstances, the median voter does not want to relinquish the ability to redistribute. Instead, she prefers a small centralization drive as increased inequality with a small reduction in aggregate capital means that prospects for redistribution are better.

4.2 Preferences for Decentralization

In the previous subsection we have shown that the same median agent that decides on taxation patterns ex post, prefers to tie her hands ex ante by voting for an interior level of decentralization. It remains to show, however, what are the preferences for decentralization of the rest of the agents in society. This turns out to be a particularly simple problem. In particular, we can write the program that the median agent solves (11), for a generic agent with type $\beta$. Denote by $\lambda(\beta)$ the solution to the following program:

$$\max_{\lambda} k(\beta) [A - \tau(\lambda)] - \frac{(k(\beta))^2}{2\beta} - \lambda + (1 - \lambda) \frac{\Phi^\gamma - 1}{\alpha}$$

(12)

The last two terms do not change due to the fact that taxation and redistribution decisions will ex post be decided by the preferred outcome of the median agent, conditional on constitutional arrangements. Hence the tension over $\lambda$ only depends on the amount of $k(\beta)$. It is intuitive, then, that the higher the ability to produce capital, the lower the level of capital taxation preferred and therefore the higher degree of decentralization favored.

**Proposition 6** Program 12 admits a unique solution, $\lambda(\beta)$. This solution is such that

$$\frac{\partial \lambda(\beta)}{\partial \beta} > 0$$

From this proposition it follows immediately that at the constitutional stage, the degree of decentralization favored by the median agent is the Condorcet Winner. Moreover, this proposition also implies that there should be a relationship between capital ownership, preferences for capital taxation and public provision and decentralization. In particular, political parties that represent capital owners should put
forward platforms that favor low taxes on capital, a smaller level of public expenditure (with respect to other political parties) and a higher degree of decentralization. In our model, decentralization becomes the way of obtaining the first two items in this agenda. The position in these dimension of such parties as the Republican Party in the US are therefore consistent with a political economy view of the degree of decentralization.

5 Heterogenous Public Goods

In this extension we show that the commitment problem prescribes a distribution of publicly provided goods between central and local administrations that is orthogonal to spillovers or differences in taste. In particular, we show that the more a publicly provided good resembles a pure transfer, the more important is the urge to decentralize its provision. The reason is that such goods will ex post exacerbate the temptation to expand public provision as redistribution.

We shall assume that the central government is constrained to provide goods uniformly across districts. We use the same utility function over publicly provided goods of the previous section, \( G(s(p)) = \frac{[s(p)]^{\alpha} - 1}{\alpha} \). However, we now assume that goods are heterogeneous. In particular, there are many types of public goods, and each is characterized by \( \alpha_n \in (0, 1] \), \( n \in N = \{1, 2, ..., N\} \) with \( \alpha_n < \alpha_{n'} \) if \( n < n' \). Furthermore, for each type of good, there is a continuum of size 1 of these type of goods.

A decentralization scheme is a set of \( \{\lambda_n\}_{n \in N} \), such that \( 0 \leq \lambda_n \leq 1 \). \( \lambda_n \) is the degree of decentralization (or localization) of goods of type \( n \): local governments will provide \( \lambda_n \) of this type of public good, while the central government will provide \( 1 - \lambda_n \). If \( g^a_m \) is the amount spent per good on goods of type \( n \) in district \( m \), and \( g^n \) is the amount on goods of type \( n \) by the central government, then the utility of agent \( i \) in district \( m \) is given by

\[
u^i (c^i, k^i, g^i, \bar{g}) = c^i_m + \sum_{n=1}^{N} \lambda_n \left( \frac{g^a_m}{X_n} \right)^{\alpha_n} - 1 + \sum_{n=1}^{N} (1 - \lambda_n) \left( \frac{g^n}{1 - \lambda_n} \right)^{\alpha_n} - 1 - \nu (k^i, \beta^i)\]

Since governments must balance their budget, total tax revenues must equal total
expenditures. Hence
\[ \sum_{n=1}^{N} g_n = \tau_m k_m + T_m \]
with an equivalent definition for the central government.

We shall now characterize the optimal decentralization schemes, i.e. all decentralization schemes that are Pareto optimal, given that players will play equilibrium strategies once the decentralization scheme is chosen.

**Proposition 7** All Pareto optimal decentralization schemes are characterized by an \( \alpha^* > 0 \), such that all public goods with \( \alpha_n < \alpha^* \) are completely centralized and all public goods with \( \alpha_n > \alpha^* \) are completely decentralized.

The intuition of the result is straightforward. Consider a decentralization scheme such that a good with a high \( \alpha' \) is centralized but one with a low \( \alpha \) is centralized. Now consider decentralizing \( \varepsilon \) of the first good, and centralizing \( \delta \) of the second, such that the equilibrium capital tax remains the same. This means that the amount of *ex post* redistribution also remains the same; note that centralizing a given amount of the first good increases the capital tax less, and hence \( \delta > \varepsilon \). Thus the only question is how efficiently the proceeds of the capital tax will be allocated. By allowing a greater number of public goods to be centralized, we are able to use the central government’s funds from the capital tax more efficiently.

In some sense, it is the goods with the most ‘redistributive power’ which are most dangerous to centralize. The reasoning is simple. Consider a public good with \( \alpha = 0 \), i.e. logarithmic. Then the tax used to provide this good is independent of the degree of inequality, and will not distort capital choices much, even if there is a large amount of inequality in society. This type of good is low in redistributive power, since its high convexity means that the return on investment beyond the efficient amount quickly drops off. On the other hand, for goods with \( \alpha \) close to 1, return on investment beyond the efficient amount drops off very slowly, so inequality will greatly distort the capital tax chosen, and hence hinder the *ex ante* investment in capital. This is counter the usual arguments that the goods that are most important to centralize are those with a high redistributational content: the greater the redistributational power, the bigger the problem of lack of precommitment.

Hence, during a constitutional stage of the game, we can see that the conflict between rich and poor should be over the extent of decentralization, not which particular goods to be decentralized. In particular, the richer the agent, the more she
benefits from increased decentralization. The intuition for this is clear from Proposition 4: as decentralization increases, the tax rate decreases, and the more capital investment one is likely to engage, the more important lowering the tax rate is.

However, even the poorest agent, one with no capital, does not want full centralization. In particular, all agents agree that goods for which $\alpha = 1$ should be fully decentralized. Otherwise, this good will be used ex post to fully redistribute wealth across society. While this may be good for the poorest agent ex post, it means that ex ante no capital investment will take place. Hence, by decentralizing some goods we provide incentives for the able to invest in capital, and some returns from this investment can be redistributed to the poor through public good provision.

Letting $\alpha^\ast (\beta)$ be the minimal $\alpha \geq 0$ that corresponds to an optimal decentralization scheme for an agent of type $\beta$, we have:

**Proposition 8** $\alpha^\ast (\beta)$ is a weakly decreasing function, and $\alpha^\ast (\tilde{\beta}) < 1$.

### 6 Conclusion

In this paper, we have proposed a theory of federalism that does not rely on assumptions about spillovers and taste heterogeneity. While these issues are important, many public goods, such as fire protection, sewers, etc. lack both significant spillover effects and substantial taste heterogeneity. It is vitally important, then, that we understand the how the centralization or decentralization of policy and budgetary decisions affects economic outcomes without these forces at work. In this view, the singular distinction between a centralized and decentralized state is the existence of constraints on policymaking due to the effects of competitive pressures. The competitive effects of federalism are many and varied, and has spawned a literature both decrying the effects and touting the virtues of such competition.\(^{16}\)

This paper contributes to that literature by considering the classic dynamic inconsistency in capital taxation problem. We have shown that federalism can be a powerful tool for a nation to precommit to certain taxation policies that it would not choose to implement ex post. This is true even in a model with benevolent governments who simply maximize their welfare.\(^{17}\) By decentralizing the provision of some public goods, the nation can effectively precommit to fund those goods using instru-

\(^{16}\)For a recent summary of the competitive effects of federalism, see McKinnon and Nechyba (1997).

\(^{17}\)If the lump-sum tax in our model were even slightly distortionary, a welfare-maximizing government would choose to use only capital taxes in our model, even if agents were completely homogenous.
ments other than capital taxes. While ex post this may require the use of inefficient tax instruments, ex ante it provides assurances to those who would choose to invest in capital generation. Hence, in a world without commitment, federalism may be a second-best solution to the problem of choosing tax policy and public investment.

References


24


7 Appendix

Proof of Lemma 1. The cross partial of the agent’s maximand is

\[
\frac{\partial^2}{\partial r \partial k^n} \left( r^k_j + F(k_j) - \rho_j k_j - T_j - T + \int_0^1 G(s_j(p)) \, dp + \int_1^1 G(s(p)) \, dp \right) = 1
\]

Hence, by Topkis’ theorem, the amount of capital that agent \( n \) invests in is weakly decreasing in the capital tax. ■

Proof of Proposition 1. The problem for each agent within the district, if he were allowed to choose policy is

\[
\max_{k_j, T_j, \tau_j} \left\{ F(k_j) - \rho_j k_j - T_j + \lambda G \left( \frac{\tau_j k_j + T_j}{\lambda} \right) \right\}
\]

by substituting in the budget constraint, subject to the constraint from capital mobility, namely that

\[
r = F'(k_j) - \tau_j - \tau
\]

Taking the first order conditions we find (and noting that \( \rho_j = F'(k_j) \))

\[
-F''(k_j) k_j + \tau_j G'(\cdot) + \mu F''(k_j) = 0
\]

\[
G'(\cdot) = 1
\]

\[
G'(\cdot) k_j - \mu = 0
\]

(Note that the second order conditions are satisfied, so we are at a maximum.) Thus we can calculate the taxes using the constraints as

\[
\tau_j = 0
\]

\[
G' \left( \frac{T_j}{\lambda} \right) = 1
\]

■

Proof of Proposition 2. Consider the problem of the voter with the median amount of capital who sets central government policy.

\[
\max_{T, \tau, s(p)} \left\{ -\tau k_{med} - T + \int_0^1 G(s(p)) \, dp \right\}
\]
subject to the budget constraint

\[ \int_{\lambda}^{1} s(p) \, dp = \tau k + T \]

and the constraint on land taxes

\[ T \geq 0 \]

Since it is immediate that spending will be equal across all national public goods, we have that the maximization problem is

\[ \max_{T, \tau} \left\{ -\tau k^{\text{med}} - T + (1 - \lambda) G \left( \frac{\tau k + T}{1 - \lambda} \right) \right\} \]

and taking the first-order conditions we have

\[-1 + G' (\cdot) = \nu \]

\[ G' (\cdot) = \frac{k^{\text{med}}}{k} = \Phi^{-1} \]

Hence, since \( \Phi^{-1} < 1 \), \( G' (\cdot) < 1 \) and the constraint on land taxes binds. So we have that

\[ T = 0 \]

\[ G' \left( \frac{\tau k}{1 - \lambda} \right) = \Phi^{-1} < 1 \]

That this is a Condorcet winner in the policy space is shown in the text. ■

**Proof of Lemma 2.** The proof that the median voter prefers a lower tax rate is given in text. ■

**Proof of Proposition 3.** Implicitly differentiating the expression for capital taxes obtained in Propostion 3, we have

\[ \frac{d}{d\lambda} \left[ \frac{\tau k}{1 - \lambda} = (G')^{-1} (\Phi^{-1}) \right] \]

\[ \frac{\partial (\tau k)}{\partial \tau} \frac{\partial \tau}{\partial \lambda} (1 - \lambda) + \tau k \]

\[ (1 - \lambda)^2 = 0 \]

\[ \frac{\partial \tau}{\partial \lambda} = -\frac{\partial (\tau k)}{\partial \tau} (1 - \lambda) < 0 \]
so long as \( \frac{\partial (r_k)}{\partial \tau} > 0 \), i.e. we are in a standard equilibrium.

**Proof of Proposition 4.** By plugging (10) in (9), obtain a second degree equation on \( k_{med} \). The largest solution to this equation is

\[
k_{med} = \frac{1}{2} \left[ A \beta_{med} + \sqrt{A^2 (\beta_{med})^2 - 4 \beta_{med}^2 \Phi_\gamma (1 - \lambda)} \right] \tag{13}
\]

This is the only solution consistent with a standard equilibrium. The comparative statics are immediate from this expression.

**Proof of Proposition 5.** The first order condition of (11) can be written as:

\[
\tau \frac{\partial k_{med}}{\partial \lambda} = \frac{\Phi_\gamma - 1}{\gamma}
\]

which, substituting in (10) can be further simplified to

\[
\Phi_\gamma \left[ \frac{1 - \lambda \frac{\partial k_{med}}{\partial \lambda}}{k_{med}} - \frac{1 - \frac{1}{\Phi_\gamma}}{\gamma} \right] = 0 \tag{14}
\]

From (13) it is easy to obtain that \( \frac{\partial k_{med}}{\partial \gamma} < 0 \), \( \frac{\partial^2 k_{med}}{\partial \lambda^2} > 0 \) and \( \frac{\partial^2 k_{med}}{\partial \gamma^2} < 0 \). The last term ensures that the second order condition for (11) holds. Totally differentiating (14) yields:

\[
\frac{d\lambda}{d\gamma} = -\frac{\frac{1 - \lambda}{k_{med}} \frac{\partial k_{med}}{\partial \lambda} + \frac{1 - \lambda}{k_{med}} \frac{\partial^2 k_{med}}{\partial \gamma^2} - \frac{1}{\Phi_\gamma}}{\frac{\partial k_{med}}{\partial \lambda} + \frac{1 - \lambda}{k_{med}} \frac{\partial^2 k_{med}}{\partial \lambda^2}}
\]

Using the signs above and the fact that \( \frac{1 - \lambda}{\Phi_\gamma} < 0 \), it is clear that the denominator is negative, while the numerator is also negative. Hence, we obtain \( \frac{d\lambda}{d\gamma} > 0 \).

Again, from (13) it follows that \( \frac{\partial k_{med}}{\partial A} > 0 \) and \( \frac{\partial^2 k_{med}}{\partial \lambda A} < 0 \). Totally differentiating with respect to productivity:

\[
\frac{d\lambda}{dA} = -\frac{\frac{1 - \lambda}{k_{med}} \frac{\partial k_{med}}{\partial \lambda} + \frac{1 - \lambda}{k_{med}} \frac{\partial^2 k_{med}}{\partial \lambda A}}{\frac{\partial k_{med}}{\partial \lambda} + \frac{1 - \lambda}{k_{med}} \frac{\partial^2 k_{med}}{\partial \lambda^2}}
\]

The denominator is again negative, while the numerator is positive. Hence, we obtain \( \frac{d\lambda}{dA} < 0 \).
Finally, for completion, the implicit function of inequality yields:

\[
\frac{d\lambda}{d\Phi} = -\frac{\partial}{\partial \Phi} \left( \frac{1}{k_{med}} \frac{\partial k_{med}}{\partial \lambda} \right) + \frac{1}{k_{med}} \frac{\partial^2 k_{med}}{\partial \lambda^2} - \frac{1}{\Phi^{1+1}}
\]

And the sign is ambiguous as \( \frac{\partial k_{med}}{\partial \Phi} < 0 \) and \( \frac{\partial^2 k_{med}}{\partial \lambda \partial A} > 0 \) but \( -\frac{1}{\Phi^{1+1}} < 0 \).

**Proof of Proposition 6.** Differentiate the objective function of program 12 with respect to \( \beta \).

\[
\frac{\partial k}{\partial \beta} \left[ A - \tau (\lambda) - \frac{k (\beta)}{\beta} \right] + \frac{(k (\beta))^2}{2\beta^2}
\]

The first term is always 0 due to capital endogenously chosen. Hence the cross-derivative is simply

\[
\frac{k (\beta)}{\beta^2} \frac{\partial k}{\partial \lambda} > 0
\]

as established in Proposition 4. By Topkis’ Theorem, the result follows.

**Proof of Proposition 7.** Consider any decentralization scheme \( \{\lambda_n\}_{n \in N} \) such that there exists a \( n', n \) such that \( n < n', \lambda_n > 0, \lambda_{n'} < 1 \). The utility of the agent is given by

\[
\frac{r k (\beta^i, r) + F (\tilde{k} (r)) - r \tilde{k} (r) - \sum_{n=1}^{N} \lambda_n + \sum_{n=1}^{N} (1 - \lambda_n) \frac{\Phi_{\gamma_n} - 1}{\alpha_n} - \nu (k (\beta^i, r), \beta^i)}
\]

Now consider another decentralization scheme \( \{\hat{\lambda}_n\}_{n \in N} \) such that

\[
\hat{\lambda}_n = \lambda_n - \delta \\
\hat{\lambda}_{n'} = \lambda_{n'} + \varepsilon \\
\hat{\lambda}_{\hat{n}} = \lambda_{\hat{n}} \text{ for all } \hat{n} \neq n, n'
\]

where

\[
\frac{\delta}{\varepsilon} = \frac{\Phi_{\gamma_{n'}}}{\Phi_{\gamma_n}}
\]

and \( \delta \) is small enough that \( \hat{\lambda}_{n'} < 1 \) and \( \hat{\lambda}_n > 0 \). Note that this change in the decentralization scheme holds the equilibrium capital tax constant, and hence holds equilibrium investment capital decisions constant. Hence, we need only calculate changes in utility due to differences in public goods and head taxes. Furthermore, the change in outcomes affects each agent in the same way, so all agents will agree on whether it is good or bad.
For public goods of type \( n \), there will be more provision, since these have now been centralized. In particular, agents will gain \( \Phi^{\gamma_n-1} \alpha_n \) in utility from these newly centralized goods. Similarly, agents will lose \( \varepsilon \frac{\Phi^{\gamma_{n'}-1}}{\alpha_{n'}} \) from the fact that less of goods of type \( n' \) will now be provided. Finally, head taxes will decrease by \( \delta - \varepsilon = (\Phi^{\gamma_{n'}-\gamma_n} - 1) \varepsilon \). Adding together these three effects and dividing by \( \varepsilon \) gives us

\[
(\Phi^{\gamma_{n'}-\gamma_n} - 1) + \Phi^{\gamma_{n'}-\gamma_n} \frac{\Phi^{\gamma_n-1}}{\alpha_n} - \frac{\Phi^{\gamma_{n'}-1}}{\alpha_{n'}} > \Phi^{\gamma_{n'}-\gamma_n} \left( \frac{\alpha_n-1}{\alpha_n} \right) + \Phi^{\gamma_{n'}} \left( \frac{1}{\alpha_n} \right) - \frac{1}{\alpha_{n'}},
\]

\[
> \Phi^{\gamma_{n'}} \left( \frac{1}{\gamma_n} - \frac{1}{\gamma_{n'}} \right) (1 - \Phi^{-\gamma_n})
\]

\[
> 0
\]

as \( \Phi > 1 \) and \( \gamma_n < \gamma_{n'} \). 

**Proof of Proposition 8.** For the first part, we wish to solve for each agent

\[
\max_{\{\lambda_n\}} \left\{ r k (\beta^i, r) + F (k (\beta^i, r)) - \sum_{n=1}^{N} \lambda_n + \sum_{n=1}^{N} (1 - \lambda_n) \frac{\Phi^{\gamma_n-1}}{\alpha_n} - \nu (k (\beta^i, r), \beta^i) \right\}
\]

It is enough to show, given the previous proposition, that for each \( \lambda_n \), the cross-partial of the objective function with respect to \( \lambda_n \) and \( \beta \) is positive, as then the most-preferred \( \lambda_n \) must be increasing in the parameter \( \beta^i \) by Topkis’ theorem.

Taking the derivative of the objective function with respect to \( \beta^i \), we have

\[
(r - \nu_1 (k (\beta^i, r), \beta^i)) k_1 (\beta^i, r) - \nu_2 (k (\beta^i, r), \beta^i) = -\nu_2 (k (\beta^i, r), \beta^i)
\]

where the equality comes from the condition that the agent is optimizing his capital choice given the rate of return, so that \( r = \nu_1 (k (\beta^i, r), \beta^i) \). Taking the derivative now with respect to \( \lambda_n \), we have

\[
-\nu_{12} (k (\beta^i, r), \beta^i) k_2 (\beta^i, r) \frac{dr}{d\lambda_n}
\]

However, \( -\nu_{12} (k (\beta^i, r), \beta^i) \) is positive by assumption, and the proof of Proposition 3 tells us that the derivative of \( k \) with respect to \( r \) is positive. Finally, Proposition 1 and 4 show that \( \frac{dr}{d\lambda_n} \) is positive, and hence by Topkis’ theorem we are done.

To see the second part, note that if \( \alpha_N < 1 \), then \( \alpha^* (\beta) < 1 \) is less than 1 by definition. Otherwise, the median voter will use this perfectly redistributive instrument and set the capital tax equal to the pre-tax return on capital *ex post*. Hence, there will be no capital investment *ex ante*. Then any agent could be better.
off if every public good was decentralized, ensuring a capital tax of zero so he not be worse off. ■
Figure 1. The graph shows possible policy alternatives to the median voter's ideal point.