

An Index of Capital Tax Competition *

Sam Bucovetsky

York University, Toronto

sam@dept.econ.yorku.ca

April 9, 2008

Abstract

If very specific assumptions are made about the production technology (output per worker is a quadratic function of the capital/labor ratio), people's preferences (identical within any jurisdiction ; linear in private good consumption and in public expenditure per capita), and capital supply (fixed), then equilibrium tax rates can be derived in closed form when jurisdictions choose their source-based tax rates on capital non-cooperatively. The focus of this paper is how the size distribution of the population of the different jurisdictions affects equilibrium tax rates. It is shown that the (population-weighted) average tax rate is determined by a (relatively) simple index, which must increase as the population distribution becomes more concentrated. The effects of tax harmonization by some subset of the jurisdictions can then be analyzed. Any tax harmonization by a group of jurisdictions must benefit residents of all jurisdictions not in the group. It must also benefit residents of the largest jurisdiction in the group, and must increase the average payoff of all residents of the group of jurisdictions doing the harmonizing.

*This paper is a (substantial) revision of the first half of "Tax Bloc Formation among Leviathans", generalizing the assumption in that earlier paper that policy makers care only about tax revenue. I thank seminar participants at the University of Warwick, the University of Illinois, Queen's University, the CEPII/CERGE workshop on "Tax Harmonization in the EU-25" and the IEB Workshop on Fiscal Federalism (Seville), as well as two anonymous reviewers and the editor, for their helpful suggestions.

1 Introduction

When jurisdictions rely on source-based taxation of a mobile factor, there are potential gains from policy coordination.¹ The simplest form of policy coordination might be an agreement by two or more jurisdictions to set a common tax rate. However, if there are differences among jurisdictions, there may be no common tax rate which makes all jurisdictions better off, without any further side payments.

Bucovetsky (1991), and Wilson (1991) showed that the smaller jurisdiction would levy a lower tax rate, and would be better off, if there were only two jurisdictions in a federation, and if these two jurisdictions differed only in population. The notion that smaller jurisdictions levy lower tax rates, and do better in tax competition, emerges as well in the well-known paper by Kanbur and Keen (1993) on commodity taxes.² Asymmetries³ between jurisdictions mean that coordination on a common tax rate may not be possible, if no side payments are allowed, even though the Nash equilibrium outcome is inefficient.

Even if jurisdictions are identical in every respect, policy coordination may induce asymmetries. If m of the n identical jurisdictions agree to set a common tax rate, then a new population distribution emerges, one with one large unit (composed of the jurisdictions which agree to coordinate), and $n - m$ small jurisdictions. In this case, while each jurisdiction may gain from tax coordination among all n jurisdictions, it may gain even more by holding out, and remaining as a small free rider in a world in which the remaining $n - 1$ jurisdictions agree to coordinate on a high common tax rate. The potential advantage to holding out should influence any negotiation among jurisdictions, whether or not side payments among them are possible.

¹The canonical model of this type of tax competition, with a large number of jurisdictions, was presented by Wilson (1986), and by Zodrow and Mieszkowski (1986). Wildasin (1988) extended the model to a finite number of “large” jurisdictions. Wilson (1999) provides an excellent survey of the literature. Many expositions of the costs and benefits of decentralization — Oates (1972), Musgrave (1997) and Sato (2003) for example — list inadequate coordination of taxes on mobile bases as one of the main drawbacks of decentralization.

²Hansen and Kessler (2001) show that the correlation between population and tax rates emerges as well when jurisdictions levy redistributive income taxes on mobile individuals.

³Hwang and Choe (1995) consider differences in capital ownership per person, as well as population.

Burbidge et al (1997) present a formal model of coalition formation in which each jurisdiction considers all the possible alternative coalition structures. In their model, individual jurisdictions choose to partition themselves into several coalitions, and these coalitions then play non-cooperatively. Applied to tax competition, this model implies several blocs of jurisdictions, each of which has agreed to a common tax rate among its members, and each of which treats each other bloc's tax rates as given. Burbidge et al (1997) consider explicitly the process of coalition formation, when all jurisdictions are aware of the non-cooperative play once coalitions have been formed. They examine when this process leads to policy coordination among all jurisdictions (formation of a "grand coalition").

To apply this sort of analysis to tax competition, we need to know what happens when blocs of countries coordinate their tax rates (and behave non-cooperatively towards other blocs). That is the main goal of this paper, to derive Nash equilibrium outcomes when tax rates are set non-cooperatively by several different governments, which may differ in their population. Due to strong assumptions about the primitives of the model, an index of tax competition can be calculated, depending only on the size distribution of the tax-setting blocs. The earlier results of two-country models are generalized — smaller blocs levy lower tax rates in equilibrium, and have higher payoffs per capita. But the consequences of tax coordination can also be derived. Any "merger", a decision by two or more formerly independent blocs to set a common tax rate, must lead to a higher average tax rate for the federation as a whole, and for higher tax rates, and higher payoffs, to each country not party to the merger. As in the two-country models, the larger party to any bilateral merger must gain from the merger (even when there are countries outside, not party to the merger), but the smaller party may lose. Finally, conditions are derived for the stability of the coalition of the whole, in which all countries agree to set a common tax rate. In a two-country world, the smaller country will gain from a common tax rate only if size differences are not too large. With a larger number of countries, it turns out that⁴ the relative size of the smallest country is again all that matters for the stability of the coalition of the whole. As long as the smallest country is not too small, complete coordination of tax rates may emerge from voluntary negotiation, even with no side payments.

⁴under some assumptions

Peralta and van Ypersele (2005) have already analyzed the Nash equilibrium of an n -jurisdiction tax competition game⁵, in which jurisdictions may differ both in size and in capital endowment per person.⁶ In their model, jurisdictions use source-based capital taxation to finance lump-sum redistribution to (identical) residents. Manipulation of the terms of trade is the incentive to use capital taxation ; capital exporting jurisdictions will subsidize capital. Here, instead, source-based capital taxation is used to finance public expenditure which cannot be provided by the private sector, and which is more valuable than private consumption, so that the marginal cost of public funds exceeds 1. The relative value of public expenditure influences the equilibrium tax rates ; the greater the marginal cost of public funds, the higher will be the equilibrium capital tax rate.

A very stylized model of tax competition is developed here, with strong assumptions on functional forms (for preferences and technology) enabling explicit solution for equilibrium tax rates.⁷ The key features of the model are

- The aggregate supply of the mobile factor is fixed.
- Each jurisdiction is populated by identical residents.
- The preferences of the residents are linear, in a numéraire private good, and in public expenditure per capita. (That is, the marginal value of an additional dollar of tax revenue is a constant.)
- Output in each jurisdiction is a quadratic function of the quantity of capital employed there.

Section 2 provides more detail on the structure of the model. The properties of the index of tax competition, are presented in section 3. In particular, it is shown there that

⁵I thank an anonymous referee for providing this reference.

⁶In Peralta and van Ypersele (2003, 2006), the authors also analyze what sort of policy coordination would be Pareto-improving in this model.

⁷Potentially, the predictions of this model could be tested, given data on different federations which differ in the size composition of their constituent jurisdictions. Brülhart and Jametti (2006), for example, use data on taxation by municipalities in different Swiss cantons to distinguish between different models of tax competition.

this index depends on the size distribution of population in the different jurisdictions. Any movement of population from one jurisdiction to another larger jurisdiction must increase the value of the index, and thus the equilibrium tax rates in all other jurisdictions (except the source and destination for the population movement). A merger of two jurisdictions into a single jurisdiction is a special case of this sort of population movement. The payoffs to the different jurisdictions (or blocs of jurisdictions) are analyzed in section 4. There the Bucovetsky–Wilson result that smaller jurisdictions do better is generalized. As well, the changes in payoffs induced by merger are considered. It is shown that the larger party to any bilateral merger must benefit, whereas the smaller party may lose. In section 5, the question addressed by Burbidge et al, the feasibility of the grand coalition, is examined, given the impact of merger on payoffs discussed in the previous section. A few concluding remarks are provided in section 6.

2 The Model

The model is one which has perhaps become standard in the analysis of source-based tax competition, with a few stronger assumptions about functional form. In one respect the model is more general than many that have been used in this literature : the number of competing jurisdictions can be arbitrary.

A federation consists of N different jurisdictions. All residents of all jurisdictions are identical. Residents are completely immobile. The preferences of a resident of jurisdiction i can be represented

$$u_i = x_i + (1 + \varepsilon)z_i \tag{1}$$

where u_i is the payoff to any resident of jurisdiction i , x_i is per capita consumption of a private numéraire good, and z_i is tax revenue per capita from a source-based capital tax. The parameter $1 + \varepsilon$ represents the marginal cost of public funds.⁸ It will be assumed that

$$\varepsilon \geq 0$$

⁸if $\varepsilon = 0$, then the model reduces to that of Peralta and van Ypersele (2005)

for each jurisdiction i , so that public expenditure of a dollar is valued at least as highly as private consumption. The formulation of preferences (1) specializes the usual assumption in the literature by specifying a constant marginal rate of substitution between private and public expenditure. This specialization is needed to derive closed-form solutions for equilibrium tax rates. However, it may not be a bad approximation if it is assumed that governments have access to several other distortionary taxes in addition to the source-based capital tax. The government chooses these other taxes optimally. If the revenue from the capital tax is a relatively small fraction of total government revenues, then changes in the revenue collected from this tax may have very little effect on the marginal cost of public funds at the government's optimum, so that utility may be treated as approximately linear in the revenue from the capital tax.⁹

Each individual in each jurisdiction supplies one unit of labor inelastically. Output is produced from this labor, and from capital, using a constant-returns-to-scale production function. If k_i is the capital-labor ratio in jurisdiction i , then output per capita $f_i(k_i)$ is defined by

$$f_i(k_i) = \left(a - \frac{b}{2}k_i\right)k_i \quad 0 \leq k_i \leq \frac{a}{b} \quad (2)$$

Notice that here the technology is assumed identical in each jurisdiction. More particularly, it is also assumed that the production function (in intensive form) is quadratic, so that the marginal product of capital is a linear function of the capital-labor ratio in the jurisdiction. Of course this is a strong assumption, again made purely for tractability.¹⁰

Capital is perfectly mobile among jurisdictions. It earns a gross return in each jurisdiction equal to its marginal product

$$f'_i(k_i) = a - bk_i$$

⁹See Keen and Lahiri (1998), for example.

¹⁰Grazzini and van Ypersele (2003), Stöwhase (2005) and Devereux, Redoano and Lockwood (forthcoming) are among the other tax competition papers which have used this assumption.

but this return is subject to a source-based unit tax t_i . Perfect mobility of capital implies that its net return be the same in each jurisdiction (in which there is non-zero employment of capital). That net return is

$$\rho = a - bk_i - t_i \quad i = 1, 2, \dots, N \quad (3)$$

where there is no jurisdiction-specific index on ρ .

The population of jurisdictions may differ. Let

s_i : share of the federation's population resident in jurisdiction i

(so that $\sum_{i=1}^N s_i = 1$). It will be convenient to use the notational convention that the largest jurisdictions are indexed by the lowest numbers :

convention : $s_1 \geq s_2 \geq \dots \geq s_N$

As in much of the literature on tax competition, each resident of region i has the same fixed endowment of capital, e .¹¹ The constraint on the capital allocation is that total capital employed in the federation cannot exceed the total fixed endowment

$$\sum_{i=1}^N s_i k_i \leq e \quad (4)$$

with strict equality if the net return to capital ρ is strictly positive.

If equation (3) holds for each jurisdiction i , and if (4) holds as an equality, they imply an allocation of capital to jurisdictions

$$k_i = e + \frac{\bar{t} - t_i}{b} \quad (5)$$

where

$$\bar{t} \equiv \sum_{i=1}^N s_i t_i$$

is the average tax rate in the federation.

¹¹Appendix 2 extends the results to the case in which each jurisdiction is homogeneous, but in which per capita capital ownership, and the marginal cost of public funds, may differ across jurisdictions.

If, for a given set of tax rates in the different jurisdictions, equation (5) defines a non-negative allocation of capital to each jurisdiction, and this allocation implies a non-negative net return ρ to investment, then the employment of capital defined by (5) is the unique allocation of capital corresponding to the tax rates.¹²

If the tax rates are too high, on average, then the allocation defined by (5) might imply a negative net return to capital. Although it has been assumed that the supply of capital is fixed, presumably capital owners do have the option of free disposal.¹³ Therefore, an alternative possible allocation is one in which the net return to capital is zero, in which some capital is withheld, so that

$$k_i = \frac{a - t_i}{b} \quad i = 1, 2, \dots, N \quad (6)$$

and $\rho = 0$. If the solutions to (6) imply capital levels such that $\sum_{i=1}^N s_i k_i \leq e$, then the allocation of capital is defined by (6), and its net return by $\rho = 0$. On the other hand, if the solutions to (6) imply capital levels such that $\sum_{i=1}^N s_i k_i > e$, then the allocation of capital is defined by (5), and its net return by

$$\rho = a - be - \bar{t} > 0 \quad (7)$$

In the remainder of the paper, only allocations of the latter type, in which capital is fully employed and earns a positive net return, will be considered. However, whether the allocation involves capital being withheld depends on the tax rates, which are chosen by the residents. Those equilibrium taxes, in turn, depend on the population distribution among jurisdictions. Subsequently, a necessary and sufficient condition, inequality (17) depending on the population distribution (and on the technology, endowments, and preference for the public good) will be derived, which ensures that the return ρ to capital is positive in equilibrium.

If the net return to capital is positive, then equation (5) implies that

$$\frac{\partial k_i}{\partial t_i} = -\frac{1 - s_i}{b} \quad (8)$$

¹²The existence and uniqueness of the equilibrium allocation of capital (k_1, k_2, \dots, k_N) , given the tax vector (t_1, t_2, \dots, t_N) , is discussed in the first appendix.

¹³see Laussel and Le Breton (1998) for a more detailed analysis of this sort of model.

As well, equation (3) implies that

$$\frac{\partial \rho}{\partial t_i} = -(1 + b \frac{\partial k_i}{\partial t_i}) = -s_i \quad (9)$$

A resident's private income is her wage income, plus the return to her endowment of capital. Constant returns to scale imply that the wage per worker w_i equals $f_i(k_i) - f'(k_i)k_i$ so that

$$w_i = \frac{1}{2}b(k_i)^2 \quad (10)$$

Since

$$x_i = w_i + \rho e$$

the tax rate t_i which maximizes the payoff $u_i = x_i + (1 + \varepsilon)t_i k_i$ to residents of jurisdiction i , taking as given the tax rates in other jurisdictions, must satisfy the first-order condition

$$\frac{\partial u_i}{\partial t_i} = bk_i \frac{\partial k_i}{\partial t_i} - (1 + b \frac{\partial k_i}{\partial t_i})e + (1 + \varepsilon)(k_i + t_i \frac{\partial k_i}{\partial t_i}) = 0 \quad (11)$$

From (8), if $\rho > 0$ then

$$\frac{\partial u_i}{\partial t_i} = s_i(k_i - e) + \varepsilon k_i - \frac{1 - s_i}{b}(1 + \varepsilon)t_i \quad (12)$$

Substituting the equation (5) defining the allocation of capital, equation (12) implies that

$$b \frac{\partial u_i}{\partial t_i} = b\varepsilon e + (\varepsilon + s_i)\bar{t} - [1 + \varepsilon(2 - s_i)]t_i \quad (13)$$

At any Nash equilibrium, jurisdiction i chooses a tax rate such that $\frac{\partial u_i}{\partial t_i} = 0$, implying that

$$t_i = \frac{1}{1 + \varepsilon(2 - s_i)}[\varepsilon(be + \bar{t}) + s_i\bar{t}] \quad (14)$$

The partial derivative of the right side of equation (14) with respect to s_i is positive, so that

Lemma 1 *In any Nash equilibrium in which $\rho > 0$, $s_i > s_j$ implies that jurisdiction i levies a higher tax rate than jurisdiction j .*

Lemma 1 is not a new result. Wilson (1991) derived this result without the restrictions imposed here on utility and production functions. However the lemma here does show that Wilson's result can be extended to federations with more than two jurisdictions.

If both sides of equation (14) are multiplied by s_i , and the resulting equation summed over all jurisdictions, the average tax rate in the federation is defined by

$$\left[1 - \sum_{i=1}^N \frac{s_i(\varepsilon + s_i)}{1 + \varepsilon(2 - s_i)} \right] \bar{t} = \left[\sum_{i=1}^N \frac{s_i}{1 + \varepsilon(2 - s_i)} \right] be\varepsilon \quad (15)$$

Equation (15) defines the equilibrium average tax rate in the federation as a function of the size distribution of the jurisdictions, of the endowments e owned in each jurisdiction, and of the marginal cost of funds $1 + \varepsilon$ in each jurisdiction. The average tax rate is a measure of the extent of tax competition.¹⁴ Most of the remainder of this paper consists of manipulation of equation (15), showing how changes in the size distribution of population affect the extent of tax competition.

3 Properties of the Index of Tax Competition

Equation (15) can be written

$$\bar{t} = \frac{\alpha be}{1 - \alpha - \beta} \quad (16)$$

where the terms α and β in the above equation are defined as

$$\alpha = \sum_{j=1}^N \frac{s_j \varepsilon}{1 + (2 - s_j) \varepsilon}$$

¹⁴Given the endowment of capital, and the technology, equation (7) shows that the net return to capital depends on the average tax rate alone, so that \bar{t} is the one measure of tax competition which would matter to capital owners. However \bar{t} is certainly not a sufficient statistic for the tax revenue per capita, $\sum s_i t_i k_i$. Equation (5) shows that a mean preserving spread in the jurisdictions' tax rates will lower aggregate tax revenue.

$$\beta = \sum_{j=1}^N \frac{(s_j)^2}{1 + (2 - s_j)\varepsilon}$$

The index of tax competition ζ is then defined as

$$\zeta \equiv \frac{\alpha}{1 - \alpha - \beta}$$

so that

$$\bar{t} = \zeta be$$

Holding constant a , b and e , the average tax rate is proportional to the index ζ , and the net return to capital is a decreasing linear function of ζ .

As well, whether the net return to capital is positive depends on ζ . From the definition (7) of the equilibrium net return to capital, ρ will be positive if and only if

$$a > be(1 + \zeta) \tag{17}$$

In particular, if all jurisdictions merged, so that $s_1 = 1$ and $s_j = 0$ for all $j > 1$, then the resulting value of ζ (for any $\varepsilon > 0$) will violate condition (17).

Not surprisingly, the more important is tax revenue to the constituent jurisdictions, the higher will be tax rates.

Proposition 1

$$\frac{\partial \zeta}{\partial \varepsilon} > 0$$

The proof of this proposition, and of most of the other results, is deferred to appendix 3.

To see the effect of the population distribution on tax rates, first let α_i and β_i denote the terms in the summations defining α and β :

$$\alpha_i \equiv \frac{s_j \varepsilon}{1 + (2 - s_j)\varepsilon}$$

$$\beta_i \equiv \frac{(s_j)^2}{1 + (2 - s_j)\varepsilon}$$

Lemma 2 *Both α_i and β_i are increasing convex functions of s_i .*

An immediate consequence of this lemma is that

Proposition 2 *If $s_i \geq s_j$, increasing s_i by δ and decreasing s_j by δ must increase the index of tax competition, for any $0 < \delta \leq s_j$.*

Proof Lemma 2 shows that increasing s_i and decreasing s_j (by the same amount) must increase both α and β : the increase in α_i (and β_i) is larger than the decrease in α_j (and β_j). Since the index of tax composition ζ equals $\alpha/(1 - \alpha - \beta)$, if α and β increase, so does ζ . •

A particular consequence of Proposition 2 is that any merger of two jurisdictions must increase the average tax rate in the federation : a merger of jurisdictions i and j ($i < j$) is equivalent to increasing s_i to $s_i + s_j$ and decreasing s_j to 0.

The measure ζ of the strength of tax competition has many of the properties of a concentration index, or a measure of inequality. For given ε , it depends only on the relative sizes of the different jurisdictions in the federation. Any changes which make the population less equal — that is, any changes which can be decomposed into a sequence of population shifts from smaller to larger jurisdictions — must raise the index.

The “least equal” population distribution is one for which $s_1 = 1$ and $s_j = 0$ for all $j > 1$. In this case $\alpha = \varepsilon/(1 + \varepsilon)$, $\beta = 1/(1 + \varepsilon)$ and $\zeta \rightarrow \infty$. This extreme result should not be surprising. With a fixed capital stock, and with a premium ε on public expenditure which is positive, a single jurisdiction would want an arbitrarily high tax rate on capital. Whenever $\zeta > \frac{a}{be} - 1$, condition (17) cannot hold, and that must be the case as $\zeta \rightarrow \infty$.

The “most equal” distribution of population (an analogue of perfect competition) would be the limiting case, as $N \rightarrow \infty$ of N identical jurisdictions. Not surprisingly, an increase in the number of identical jurisdictions tends to increase tax competition.

Proposition 3 *If the world consisted of N equal-sized jurisdictions, then an increase in N must increase tax competition.*

Proof If all jurisdictions were the same size, then $s_i = 1/N$ for all i , and

$$\zeta = \frac{N}{N-1} \frac{\varepsilon}{\varepsilon+1} \quad (18)$$

which is a decreasing function of N . •

In the limit, as $N \rightarrow \infty$, with a large number of identical jurisdictions, ζ approaches its lowest possible value

$$\zeta_0 = \frac{\varepsilon}{1+\varepsilon}$$

As was clear from the analysis of Wilson (1986), and of Zodrow and Mieszkowski (1986), “perfect competition” here does not drive the tax rate down to zero, but to the positive level chosen by a government which ignores the effects of its own tax policy on the world net rate of return.

4 Who Gains?

Wilson (1991) showed that, in a two-jurisdiction world, not only do smaller jurisdictions levy lower tax rates, they also do better.

That result can be generalized to a world of more than two jurisdictions. Using the normalization that lower-numbered jurisdictions are larger :

Proposition 4 *If jurisdictions differ only in population, then $u_i \leq u_{i+1}$ in equilibrium, with equality only if the two jurisdictions have the same population.*

When jurisdictions differ in size, there are capital flows in equilibrium. Larger jurisdictions are net exporters of capital, since their tax rates are higher than average. There is a terms-of-trade effect which may work in the opposite direction from the standard tax competition externality. Although an increase in one jurisdiction’s tax rate increases other jurisdictions’ tax bases, it also reduces the net return to capital, which will tend to harm capital exporters.

If the differences among jurisdictions' tax rates are due solely to size differences, then it turns out the terms of trade effects cannot be more important than the tax base externality, as the following lemma shows.

Lemma 3 *If countries differ only in size, then if $i \neq j$, residents of country i gain from a tax increase in country j .*

In other words, the basic insight of the Wilson–Zodrow–Mieszkowski model, that tax rates must be too low if jurisdictions are identical, or if jurisdictions are all infinitesimal in size, continues to hold when size differences ¹⁵ exist.

An implication of Lemma 3 is that increases in ζ tend to be beneficial. In particular, consider a merger of any two jurisdictions. The merger replaces jurisdictions i and j with a new jurisdiction of size $s_i + s_j$ (and leaves unchanged the size of every other jurisdiction).

Proposition 5 *If two jurisdictions merge, the payoff to every other jurisdiction (other than the merging jurisdictions) must increase.*

Proof From proposition 2, a merger of jurisdictions i and j must increase ζ , and thus increase \bar{t} . Equation (14) then shows that the tax rate in every jurisdiction (other than i and j) must increase. The new tax rate in the merged jurisdiction $i \cup j$ must be higher than either jurisdiction's previous tax rate : the new merged jurisdiction has a higher total population, and the average tax rate \bar{t} has increased.

So if jurisdiction k ($k \neq i, j$) were to leave its own tax rate unchanged after the merger, Lemma 3 shows that the payoff to its residents would increase. In fact, the jurisdiction would not choose to leave its own tax rate unchanged. Therefore, its payoff in equilibrium, when it chooses a tax rate which is preferred strictly to the previous t_k , must be higher yet. This revealed preference argument shows that a merger of jurisdictions i and j must raise u_k for any $k \neq i, j$. •

Proposition 5 shows that, under the hypotheses of this model, no jurisdiction would oppose tax harmonization by other groups of jurisdictions. Instead, outsiders are major beneficiaries.

¹⁵or differences in the marginal cost of public funds ; the proof of lemma 3 would still work if ε_i were jurisdiction-specific

A similar result can hold in industrial organization. In a simple Cournot model, with identical firms producing a homogeneous good under constant returns to scale, and facing a linear demand curve, a merger of any 2 firms must increase the profits of the remaining $N - 2$ firms.¹⁶ But the parties to the merger actually lose profits by merging, in this simple Cournot model.

That is not the case in tax competition among jurisdictions. If two jurisdictions merge (with the merged entity playing as a single player in the new $N - 1$ -jurisdiction non-cooperative tax-setting game), the larger of the two jurisdictions must be better off in the new equilibrium.

Proposition 6 *If jurisdiction i merges with jurisdiction j , and $i < j$, then jurisdiction i must gain from the merger.*

Proof Here too, the result can be shown using a revealed preference argument. The merger increases ζ , and so increases the tax rates t_k in every other jurisdiction $k \neq i, j$. One feasible policy for the new merged entity would be to set its tax rate at the old (pre-merger) value of t_i . Residents of jurisdiction i must be better off in such a situation : their own tax rate is unchanged, the former jurisdiction j has increased its tax rate from t_j to $t_i \geq t_j$, and every other jurisdiction $k \neq i, j$ has increased its tax rate. Lemma 3 implies u_i must increase in such a situation.

In the new equilibrium, residents of the merged jurisdiction do not choose to stay with the old t_i as their tax rate ; they increase it. That is because the new, higher $t_{i \cup j}$ yields an even higher payoff. Therefore, $u_{i \cup j}$ in the new equilibrium must be higher than u_i in the old, pre-merger equilibrium.

This argument does not apply to jurisdiction j , the smaller of the two jurisdictions. Setting the merged entity's tax rate at t_i would involve changing jurisdiction j 's own tax rate (from t_i to t_j) and the direct negative effect of moving away from its (initial) best response may outweigh the positive effect of other jurisdictions raising their tax rates. Setting the merged entity's tax rate at t_j would mean that jurisdiction i 's tax rate has been decreased, not increased, so that Lemma 3 does not apply. •

•

¹⁶see Salant et al (1983) for this result, and Perry and Porter (1985) for analysis of its robustness.

It is only the larger party which must always gain from the merger. The smaller jurisdiction may gain, but it may lose. While the merger results in higher tax rates elsewhere, a move which benefits both merging parties, it also results in the residents of the smaller party becoming part of a bigger jurisdiction. And bigger jurisdictions do worse in any equilibrium.¹⁷

For example, suppose that there are three jurisdictions, with $s_1 = 0.925$, $s_2 = 0.04$ and $s_3 = 0.035$. If $\varepsilon = 2$, then jurisdiction 3 would lose from a merger of jurisdictions 1 and 3. However, it would gain if it merged instead with jurisdiction 2. The table below shows the tax rates, and the payoffs, for the original 3-jurisdiction example, and after each of the 3 pairwise mergers¹⁸.

Table 1

structure	taxes			payoffs			ζ	ρ
	1	2	3	1	2	3		
all 3 separate	0.308	0.144	0.143	1.427	2.776	2.775	5.61	0.705
1 and 3 harmonize	0.566	0.251	0.566	2.011	6.984	2.011	10.52	0.446
2 and 3 harmonize	0.311	0.149	0.149	1.438	2.810	2.810	5.67	0.701
1 and 2 harmonize	0.644	0.644	0.282	2.187	2.187	8.710	12.00	0.368

While it certainly is possible that the smaller jurisdiction loses from a merger, it is true that the average payoff in the merged jurisdictions must be increased by the merger.

Proposition 7 *If jurisdictions i and j merge, the average payoff $\frac{1}{s_i+s_j}(s_i u_i + s_j u_j)$ of residents of the two jurisdictions must increase.*

So a merger of jurisdictions i and j (with $i < j$) may not represent a Pareto improvement : residents of jurisdiction j may lose. But it must increase the average

¹⁷Of course, residents of the larger party also see their jurisdiction's size increase. But Proposition 6 shows that the negative effects of these size gains must be offset by the other positive effects, for the larger party.

¹⁸when $a = 20/19$, $b = 1/19$ and $e = 1$; with these parameters, the return to capital would be 1 if there were no taxes, so that the unit taxes t_i are approximately equal to the proportional tax rates

payoff in jurisdictions i and j , and therefore, from Proposition 5, the merger must increase the total payoff of all residents added together. This statement – that merger must increase the aggregate payoff — is a stronger statement than saying that merger must increase \bar{t} , since aggregate welfare does not depend just on the average tax rate, but on the variation of tax rates among jurisdictions.

5 Stability of Cooperation

Even if side payments are not possible, policy coordination among jurisdictions may benefit each participant. Suppose that the only way in which jurisdictions can coordinate policy is to form blocs, in which taxes are fully harmonized within blocs, and in which there are no side payments among the jurisdictions in a bloc. Given the potential gains from coordination, what blocs might form?

To answer that question, some notion is required of a stable coalition structure. I define a coalition structure as any partition of the N jurisdictions. That is, a coalition structure is any collection (S_1, S_2, \dots, S_M) of subsets of $\{1, 2, \dots, N - 1, N\}$, with $1 \leq M \leq N$, such that

$$S_i \cap S_j = \emptyset \quad i \neq j$$

$$\bigcup_{j=1}^M S_j = \{1, 2, \dots, N - 1, N\}$$

The M coalitions then behave as individual jurisdictions, in a tax-setting game, in which the share of the population of coalition j is simply the sum of the shares of the population of its constituent members. The results of the previous sections yield the payoff to each coalition in a structure, as a function of the size distribution of all the coalitions in the structure.

There are several notions of what constitutes a stable coalition structure in this sort of environment, depending on the possible rules for formation and dissolution of coalitions.¹⁹ Here I do not attempt to categorize the possible stable structures. Instead, only one structure is examined, the “grand coalition”, in which all jurisdictions agree to

¹⁹See Hart and Kurz (1983), Bloch (1996), Ray and Vohra (1997) for example

a common tax rate. Burbidge et al (1997) address a similar question, in a multi-stage game, in which the N jurisdictions first negotiate formation of a coalition structure, and in which the resulting coalitions then play non-cooperatively.

Here, rather than model the negotiation process at all, I derive a sufficient condition for the stability of this grand coalition : that it yield a higher payoff to any jurisdiction than that jurisdiction would get in any other coalition structure. This condition might be stronger than necessary : jurisdiction i might get a higher payoff from some other coalition structure than from the grand coalition, but might not be able to threaten credibly to put together this blocking coalition structure.²⁰ But if the condition does hold, tax competition might be a problem relatively easily solved. Without any coercion from a higher level of government, individual governments would all benefit from an agreement to harmonize tax rates.

Given the assumptions of the model here, the biggest threat to the grand coalition is easy to identify. It is the coalition structure in which the largest $N - 1$ jurisdictions form a single bloc, and the smallest jurisdiction, jurisdiction N , remains as a singleton outside that bloc. If jurisdiction N prefers the grand coalition to the coalition structure just described, then any jurisdiction will prefer the grand coalition to any other coalition structure. Proposition 8 below establishes this result, and the conditions under which this most likely threat to the grand coalition would not be credible. In deriving this result, two further assumptions are used, both of which will hold if the technology is sufficiently productive. Assumption (A1) ensures that the grand coalition chooses simply to expropriate the rents from capital. If the assumption did not hold, governments might best exercise their monopoly power by using less than the full quantity of capital available.²¹ Assumption (A2) implies that some competition still exists when there is more than one bloc.²²

²⁰For example, the sufficient condition does not hold in the example of the previous section. Jurisdiction 3 does better if it is on its own, with jurisdictions 1 and 2 harmonizing taxes, than if it joins the grand coalition. But jurisdictions 1 and 2 would not be willing to form such a coalition if jurisdiction 3 were not to join.

²¹That is, this assumption ensures that demand for capital is inelastic, so that rents would not be maximized by withholding some of it from production

²²This assumption guarantees that the equilibrium return to capital, net of taxes, is positive in any coalition structure except for the grand coalition.

$$A1 : a > be^{\frac{1+2\varepsilon}{1+\varepsilon}}$$

$$A2 : \frac{3s_N(1-s_N)}{(2-s_N)(1+s_N)}a > be$$

Proposition 8 *Given the tastes and technology, there is some minimum population share $s^0 > 0$ so that the grand coalition Pareto dominates any other coalition structure if and only if the population share s_N of the smallest jurisdiction is s^0 or greater.*

Even without an explicit model of coalition formation Proposition 8 may suggest what could be the biggest potential obstacle to tax harmonization. In this model, it is free riding by the small jurisdictions which set low source-based taxes, and benefit from high tax rates elsewhere, caused by policy coordination among the larger countries.

6 Concluding Remarks

Under strong assumptions, a potential method has been provided to compare the extent of tax competition across different federations, or across time periods. The index ζ , which depends only on the population distribution of jurisdictions within the federation, determines the average tax rate in the federation. This index shares many properties with concentration indices in industrial organization, or inequality indices. The index must increase if there is a shift of population from a small jurisdiction to a large one or (a fortiori) if jurisdictions merge.

The properties of this index also imply what are the serious obstacles to policy coordination, and what are not, if tax competition is viewed as a concern. Free riding by small countries on any policy coordination by larger countries is a problem here. But segmentation into a few large blocs is not. In contrast to some models of trade policy coordination, division of the federation into a few distinct blocs is not a likely outcome, and not an impediment to further coordination. Tax coordination among any group of jurisdictions must benefit the remaining jurisdictions. As well, it can be shown that tax coordination by one group of jurisdictions makes further coordination more attractive, at least when the public output is valued highly.

The applicability of the model presented here should not be oversold. The model predicts that small countries will levy low taxes. But the tiny, island-nation tax havens

which so vex the OECD do not really fit this model very well. Here investment in any country requires labor, which is immobile among countries. While very small countries benefit from tax competition in this model, the total amount of the tax base which gets allocated to these countries is severely constrained by the scarcity of labor. The tax havens receiving the OECD's attention tend to specialize in attracting very footloose capital, which has minimal requirement of complementary fixed inputs. A large fraction of some of these assets (oil tankers, the capital of certain financial intermediaries) winds up locating in tax havens, an outcome which is not possible with the less footloose industries considered in this paper.

Finally, there is the implication of the model which is not consistent with observed behavior, at least as usually interpreted. Even without side payments, tax coordination is often mutually beneficial for jurisdictions, and must be so if they are of similar size. Proposition 8 implies that free riding by small countries may persist, and may have a significant impact on residents of larger countries. But continued, serious tax competition among large countries of similar size is hard to reconcile with the results presented here. Unless commitment problems are so severe that a one-shot non-cooperative game accurately describes the behavior of sovereign jurisdictions, large countries can gain from policy coordination, and will gain from policy coordination by other jurisdictions.

7 Appendix 1 : Existence and Uniqueness of Equilibrium

Lemma 4 *If $a > 2b\bar{e}$, there can be no Nash equilibrium in which $\sum s_i k_i < e$.*

Proof When there is excess supply of capital, then k_i is defined by equation (6), so that

$$\frac{\partial k_i}{\partial t_i} = -\frac{1}{b} \quad (19)$$

The first-order condition (11) for each jurisdiction then becomes

$$\varepsilon_i k_i - (1 + \varepsilon_i) \frac{t_i}{b} = 0 \quad (20)$$

If t_i is optimal for jurisdiction i , and if the tax rates (t_1, t_2, \dots, t_N) lead to an excess supply of capital, then equations (6) and (20) imply that

$$t_i = \frac{a\varepsilon_i}{1 + 2\varepsilon_i} \quad (21)$$

Since the right side of (21) is an increasing function of ε_i , therefore, for each jurisdiction i ,

$$t_i < \frac{a}{2}$$

if there is excess supply of capital in equilibrium.

At least one jurisdiction must have a supply of capital k_i which is no greater than the average endowment per capita \bar{e} . The net return to capital in that jurisdiction is

$$\rho_i \equiv a - bk_i - t_i \geq a - b\bar{e} - t_i \geq \frac{1}{2}(a - 2b\bar{e})$$

which implies that the net return to capital in at least one jurisdiction must be strictly positive if $a > 2b\bar{e}$. •

Lemma 5 *There can be at most one Nash equilibrium in which the return to capital ρ is positive.*

Proof If the return to capital is positive in some Nash equilibrium, then the average tax rate \bar{t} is defined by equation (15). Since the coefficient of \bar{t} on the left side of that equation must be strictly positive when there are two or more jurisdictions, therefore (15) defines \bar{t} uniquely. But then the tax rate in each jurisdiction i is defined uniquely by equation (14). •

Lemma 6 *If $a > 2b\bar{e}$, if there is a Nash equilibrium with a positive net return to capital, then there is no other Nash equilibrium.*

Proof Suppose that t is a vector of Nash equilibrium tax rates. From lemma 5, the only other possible Nash equilibrium would be some tax vector t' which resulted in a net return to capital $\rho' = 0$, and which resulted in no excess supply of capital. When there is no excess supply of capital

$$\rho = a - b\bar{e} - \bar{t}$$

so that the fact that $\rho' < \rho$ means that $\bar{t}' > \bar{t}$.

Now let jurisdiction j be a jurisdiction for which

$$t'_j - \bar{t}' \geq t_j - \bar{t}$$

The fact that \bar{t}' and \bar{t} are weighted averages of tax rates imply such a j must exist. Since $\bar{t}' > \bar{t}$, therefore $t'_j > t_j$.

Since $k_i = \bar{e} + \frac{1}{b}[\bar{t} - t_i]$, therefore $k'_j \leq k_j$.

Now the right side of equation (12) must equal 0 when jurisdiction j sets a tax rate of t_j (and when its capital supply is k_j). Since $k'_j < k_j$, and $t'_j > t_j$, the right side of (12) must be strictly negative when the return to capital is 0, and when jurisdiction j sets a tax rate of t'_j . But, when the return to capital is $\rho' = 0$, then a slight lowering of t'_j would result in a positive return to capital, so that (12) does describe the change in the payoff to the jurisdiction from lowering its tax rate slightly from t'_j . Since this derivative is negative, lowering its tax rate slightly would raise the payoff to jurisdiction j , contradicting the assumption that t'_j is its optimal tax rate. •

Lemma 7 *If $a > 2b\bar{e}$, there exists at least one Nash equilibrium. In any Nash equilibrium, capital is fully allocated, so that $\sum_{i=1}^N s_i k_i = \bar{e}$.*

Proof The second part of the lemma follows immediately from lemma 4.

To prove existence, note that each jurisdiction's payoff u_i is a continuous function of any other jurisdiction's tax rate.

Also, differentiation of equation (13) with respect to t_i shows that

$$\frac{\partial^2 u_i}{\partial t_i^2} = -(1 - s_i)(1 + s_i + \varepsilon_i) < 0$$

when $\rho > 0$.

When there is excess supply of capital,

$$\frac{\partial u_i}{\partial t_i} = \frac{1}{b}(\varepsilon_i a - (1 + 2\varepsilon_i)t_i)$$

so that a jurisdiction's payoff is again a concave function of its own tax rate.

At the boundary between regimes, when tax rates are set so that $\rho = 0$ and $\sum_i k_i = \bar{e}$, then $\frac{\partial k_i}{\partial t_i}$ increases discontinuously in absolute value as t_i is increased, from $(1 - s_i)/b$ to $1/b$. That means that $\frac{\partial u_i}{\partial t_i}$ must fall as t_i is increased, and the allocation moves from the positive-return regime to the excess-supply regime.

Therefore, each jurisdiction's payoff is a concave function of its own tax rate.

Each jurisdiction's choice of tax rates can be bounded. Equation (12) shows that $\frac{\partial u_i}{\partial t_i} > 0$ whenever $t_i < -\frac{s_i}{1-s_i} b e_i$, if the return to capital were positive — and $\frac{\partial u_i}{\partial t_i}$ must be positive if $t_i < 0$ and if there is excess supply of capital. If $t_i > (a - b\bar{e})/s_i$ then there must be excess supply of capital, and if there is excess supply of capital, then $\frac{\partial u_i}{\partial t_i}$ must be negative if $t_i > \frac{a\varepsilon_i}{1+2\varepsilon_i}$. Therefore, the tax rates chosen by each jurisdiction can be restricted to the interval $[T_i^1, T_i^2]$ where

$$T_i^1 = -\frac{s_i}{1-s_i} b e_i$$

$$T_i^2 = \max\left((a - b\bar{e})/s_i, \frac{a\varepsilon_i}{1+2\varepsilon_i}\right)$$

The existence result (Theorem 1) of Dasgupta and Maskin (1986) can then be applied : if each player's strategy set is a compact set, if each player's payoff is a concave function of her own strategic variable, and if each player's payoff is a continuous function of the other players' strategic variables, then there must be a Nash equilibrium in pure strategies. •

8 Appendix 2 : Differences in Endowments and Tastes

Suppose now that each jurisdiction is still homogeneous, but that the endowment of capital per person e_i and the marginal cost of public funds $1 + \varepsilon_i$ may vary across jurisdictions. Let \bar{e} and $\bar{\varepsilon}$ represent the average value (weighted by population) of these parameters in the population. Allowing for variation in e_i and ε_i implies that equation (14) becomes

$$t_i = \frac{1}{1 + \varepsilon_i(2 - s_i)} [\varepsilon_i(b\bar{e} + \bar{t}) + s_i(b(\bar{e} - e_i) + \bar{t})] \quad (22)$$

This equation shows how tax rates vary across jurisdictions, in response to differences in factor ownership or in tastes. Other things equal, the lower a jurisdiction's capital endowment e_i , the higher its tax rate. This tendency should not be surprising : tax increases serve to lower the net return ρ to capital, and the bigger is e_i the more a jurisdiction's residents will suffer from such a decrease. Notice as well that as $\varepsilon_i \rightarrow \infty$, or as $s_i \rightarrow 0$, this effect of endowments on tax rates disappears. When ε_i is very high, the effect of tax rates on residents' private consumption becomes unimportant. When s_i is very low, the influence the jurisdiction has on the federation's net return to capital becomes trivial.

Further, the derivative of the right side of (14) with respect to ε_i is

$$\frac{1}{[1 + \varepsilon_i(2 - s_i)]^2} [(b\bar{e} + \bar{t})(1 - s_i(2 - s_i)) + bs_i e_i(2 - s_i)]$$

which is strictly positive whenever $s_i < 1$, so that jurisdictions which value the public sector more highly will levy higher capital tax rates in equilibrium, other things held equal.

Multiplying both sides of equation (22) by s_i , and summing over all jurisdictions, yields the following generalization of equation (15).

$$\left[1 - \sum_{i=1}^N \frac{s_i(\varepsilon_i + s_i)}{1 + \varepsilon_i(2 - s_i)} \right] \bar{t} = \left[\sum_{i=1}^N \frac{s_i \varepsilon_i}{1 + \varepsilon_i(2 - s_i)} \right] b\bar{e} + \left[\sum_{i=1}^N \frac{(s_i)^2(\bar{e} - e_i)}{1 + \varepsilon_i(2 - s_i)} \right] b \quad (23)$$

The second term on the right side of equation (23) shows the impact of distribution of capital endowment across jurisdictions on the federation's average tax rate.

Proposition 9 *Holding constant the average capital endowment \bar{e} in the federation, other things equal the average tax rate is increased by*

i shifting capital endowments from large jurisdictions to small jurisdictions

ii shifting capital endowments from jurisdictions with a low marginal cost of public funds to those with a high marginal cost of public funds

Proof First, note that the term multiplying \bar{t} on the left hand side of (15) must be strictly positive, if there is more than one jurisdiction in the federation. Since $2 - s_i > 1$ if $s_i < 1$, then $1 + \varepsilon_i(2 - s_i) > 1 + \varepsilon_i > s_i + \varepsilon_i$, so that $\sum_{i=1}^N s_i \frac{s_i + \varepsilon_i}{1 + \varepsilon_i(2 - s_i)} < 1$, implying that

$$1 - \sum_{i=1}^N \frac{s_i(\varepsilon_i + s_i)}{1 + \varepsilon_i(2 - s_i)} > 0$$

This positivity implies that a shift in endowments will raise the average tax rate \bar{t} if and only if it increases the right side of equation (15).

Suppose that the capital endowment in jurisdiction j falls by $\Delta > 0$, and that the capital endowment in jurisdiction i increases by $(s_j/s_i)\Delta$, so that the average capital endowment \bar{e} is unchanged.

From equation (15), the change in the equilibrium average tax rate in the jurisdiction will be proportional to

$$s_j \Delta \left[\frac{s_j^2}{1 + \varepsilon_j(2 - s_j)} - \frac{s_i^2}{1 + \varepsilon_i(2 - s_i)} \right] \quad (24)$$

Since

$$\frac{s^2}{1 + \varepsilon(2 - s)}$$

is an increasing function of s , and a decreasing function of ε , expression (24) will be positive if either (i) $\varepsilon_i = \varepsilon_j$ and $s_i < s_j$, or (ii) $s_i = s_j$ and $\varepsilon_i > \varepsilon_j$. •

The logic underlying part *i* of Proposition 9 was described earlier. High capital ownership in a jurisdiction reduces the incentive to increase taxes, since tax increases

lower net capital ownership income. But this effect depends on the influence one jurisdiction's unilateral tax increases have on the federation's net return to capital. This influence is proportional to a jurisdiction's size. Therefore, it is the capital ownership in the largest jurisdictions which has a disproportionate influence on the equilibrium tax rate.

The effects of changing the distribution of the marginal valuation $1 + \varepsilon_j$ of the public output are not so easy to sign. However, there is some tendency for taxes to be higher if the large jurisdictions have higher valuations.

Proposition 10 *If all jurisdictions have the same endowment e_i of capital, and if all jurisdictions initially have the same marginal valuation of public output $1 + \varepsilon_i$, then the equilibrium average tax rate \bar{t} will increase if ε_i increases by a small amount in some large jurisdiction, and falls in some smaller jurisdictions, so as to keep constant the average valuation $\bar{\varepsilon} \equiv \sum s_i \varepsilon_i$.*

Proof

Suppose that ε_j falls by Δ , and ε_i increases by $(s_j/s_i)\Delta$, so that $\bar{\varepsilon}$ is unchanged.

Equation (15) implies that

$$\left[1 - \sum_{i=1}^N \frac{s_i(\varepsilon_i + s_i)}{1 + \varepsilon_i(2 - s_i)} \right] \frac{\partial \bar{t}}{\partial \Delta} = s_j(b\bar{\varepsilon} + \bar{t}) \left[\frac{1}{(1 + \varepsilon(2 - s_j))^2} - \frac{1}{(1 + \varepsilon(2 - s_i))^2} \right] + s_j \bar{t} \left[\frac{s_j(2 - s_j)}{(1 + \varepsilon(2 - s_j))^2} - \frac{s_i(2 - s_i)}{(1 + \varepsilon(2 - s_i))^2} \right] \quad (25)$$

when $\varepsilon_i = \varepsilon_j \equiv \varepsilon$.

Since

$$\frac{1}{(1 + \varepsilon(2 - s))^2}$$

is an increasing function of s , and

$$\frac{s(2 - s)}{(1 + \varepsilon(2 - s))^2}$$

is a decreasing function of s , each term on the right side of equation (25) will be positive if and only if $s_i > s_j$.

•

9 Appendix 3 : (Most of the) Proofs

Proof of Proposition 1 Increasing ε will decrease β . To show the result, it is sufficient to demonstrate that $\alpha + \beta$ must increase with ε .

Since

$$\alpha + \beta = \sum_{j=1}^N \frac{s_j(\varepsilon + s_j)}{1 + (2 - s_j)\varepsilon}$$

it is sufficient to show that

$$\frac{s_j(\varepsilon + s_j)}{1 + (2 - s_j)\varepsilon}$$

is an increasing function of ε . The derivative of this expression with respect to ε is proportional to

$$s_j(1 + \varepsilon(2 - s_j)) - (2 - s_j)s_j(\varepsilon + s_j) = s_j[1 - s_j(2 - s_j)]$$

The term $s_j(2 - s_j)$ is an increasing function of s_j when $0 < s_j < 1$, and it equals 1 when $s_j = 1$, so that $s_j(2 - s_j) < 1$ for all $0 \leq s_j < 1$, proving that the term in square brackets in the equation above must be positive. •

Proof of Lemma 2

$$\frac{\partial \alpha_j}{\partial s_j} = \frac{\varepsilon(1 + 2\varepsilon)}{[1 + (2 - s_j)\varepsilon]^2}$$

and

$$\frac{\partial \beta_j}{\partial s_j} = \frac{2s_j(1 + 2\varepsilon)}{[1 + (2 - s_j)\varepsilon]^2}$$

•

Proof of Proposition 4 When jurisdictions differ only in size, the only variable which matters for residents' payoffs, and which differs in equilibrium across jurisdictions, is the jurisdiction's tax rate. So to compare equilibrium payoffs across jurisdictions, the effects of differing tax rates must be considered, holding the average tax rate at the true equilibrium tax rate \bar{t} .

Equation (14) shows how a jurisdiction's tax rate varies with its size s_i . Let $t(s)$ denote the solution to equation (14), as a function of a jurisdiction's size — holding constant not only the marginal cost of public funds ε and the endowment e , but also the average tax rate \bar{t} . Lemma 1 shows that $t'(s) > 0$.

So in any equilibrium in which the size distribution of jurisdictions is (s_1, s_2, \dots, s_N) , then the tax rate t_i is just $t(s_i)$, if the average tax rate \bar{t} is set correctly (to the rate defined by equation (15)).

Then let $k(s)$ be the capital supplied to a jurisdiction when its tax rate is $t(s)$, again holding constant the average tax rate \bar{t} . Equation (5) shows that

$$k'(s) = -\frac{1}{b}t'(s)$$

In an equilibrium, the payoff to a resident of a jurisdiction of size s is

$$u(s) \equiv f[k(s)] - f'[k(s)]k(s) + \rho e + (1 + \varepsilon)t(s)k(s)$$

so that

$$u'(s) = \varepsilon t'(s)k(s) + (1 + \varepsilon)t(s)k'(s)$$

which means that

$$u'(s) = \left[\varepsilon k(s) - \frac{1 + \varepsilon}{b}t(s) \right] t'(s) \tag{26}$$

From the first-order condition for optimality (12),

$$\varepsilon k_i - \frac{1 + \varepsilon}{b}t_i = s_i \left[e - k_i - \frac{1 + \varepsilon}{b}t_i \right]$$

so that the definition (5) of k_i implies that

$$\varepsilon k_i - \frac{1 + \varepsilon}{b}t_i = -\frac{s_i}{b}[\bar{t} + \varepsilon t_i]$$

All tax rates must be positive (when jurisdictions have the same endowments and cost of capital) — this follows from equations (16) and (14).

Therefore, it must be the case that $\varepsilon k_i - \frac{1 + \varepsilon}{b}t_i < 0$, so that expression (26) must be negative.

Since the payoff of a resident in jurisdiction i is $u(s_i)$, residents of smaller jurisdictions must have higher payoffs in equilibrium. •

Proof of Lemma 3 The effect of t_j on country i 's payoff is

$$\frac{\partial u_i}{\partial t_j} = (e - k_i) \frac{\partial \rho}{\partial t_j} + (1 + \varepsilon) t_i \frac{\partial k_i}{\partial t_j} \quad (27)$$

Since $\frac{\partial \rho}{\partial t_j} = -s_j$, and $\frac{\partial k_i}{\partial t_j} = \frac{s_j}{b}$, therefore expression (27) is proportional to

$$(1 + \varepsilon) t_i - b(e - k_i)$$

From equation (11),

$$-\frac{\partial k_i}{\partial t_i} [(1 + \varepsilon) t_i - b(e - k_i)] = (1 + \varepsilon) k_i - e$$

so that to prove the result, it is sufficient to show that $(1 + \varepsilon) k_i > e$.

But equation (12) can also be written

$$(1 + \varepsilon) k_i - e = (1 - s_i) [(k_i - e) + \frac{1 + \varepsilon}{b} t_i]$$

Since

$$k_i - e = \frac{\bar{t} - t_i}{b}$$

therefore

$$(1 + \varepsilon) k_i - e = \frac{1 - s_i}{b} [\bar{t} + \varepsilon t_i] \quad (28)$$

When endowments are identical in all jurisdictions, equations (15) and (14) imply that all tax rates are positive, so that equation (28) shows that $\frac{\partial u_i}{\partial t_j} > 0$. •

Proof of Proposition 7 Suppose that tax rates stayed constant in jurisdictions other than i and j , and that jurisdictions i and j agreed to set a harmonized common tax rate, equal to the average of their previous tax rates

$$t' = \frac{s_i}{s_i + s_j} t_i + \frac{s_j}{s_i + s_j} t_j$$

Since tax rates elsewhere are unchanged, and since the tax rate in jurisdictions i and j combined is unchanged, the overall allocation of capital $s_i k_i + s_j k_j$ to the two jurisdictions is the same. The payoff to a resident of region i can be written

$$u_i = f(k_i) + \rho(e - k_i) + \varepsilon t_i k_i$$

so that

$$s_i u_i + s_j u_j = \rho(e - [s_i k_i + s_j k_j]) + s_i f(k_i) + s_j f(k_j) + \varepsilon(s_i t_i k_i + s_j t_j k_j)$$

The first term in the above expression is unchanged if t' is the average tax in the 2 jurisdictions ; $f(k_i)$ and $t_i k_i$ are both concave functions of t_i , so that the average payoff in the two jurisdictions must increase.

Merger does not leave the tax rates the same in other jurisdictions ; it increases them. Lemma 3 shows that the effect of these outside tax increases is to increase the average payoff further.

The merged jurisdictions do not choose to levy a tax rate of t' ; but they do better still when they choose their payoff-maximizing common tax rate. •

Proof of Proposition 8 For concreteness, denote as the “deviating coalition structure” the coalition structure in which jurisdiction N is alone, and the rest of the jurisdictions form a single coalition, that is for which

$$S_1 = \{1, 2, \dots, N - 1\}$$

$$S_2 = \{N\}$$

Assumption *A2* ensures that the return to capital ρ is positive in the deviating coalition structure, as is shown in Lemma 8 below. Proposition 2 implies that the deviating coalition structure has the highest value for ζ of any coalition structure, except for the grand coalition. Therefore, if assumption *A2* holds, then the net return to capital is positive in the tax-setting equilibrium for any coalition structure, except for the grand coalition, so that the analysis of the previous sections applies.

Given the fixed supply of capital, the grand coalition would want to raise the uniform tax t on capital in the federation at least up to the point at which the net return to capital has been driven to zero : if $\rho > 0$, then any increase in t simply converts private consumption to public consumption, which is worthwhile if $\varepsilon \geq 0$. A tax rate of $t = a - be$ just drives the return to capital down to zero. The grand coalition might consider increasing t still higher, and forcing the withdrawal of some capital. Assumption *A1* ensures that this further increase is not worthwhile, as Lemma 9 below shows. Given that $t = a - be$, then $z = (a - be)e$ and $x = \frac{1}{2}be^2$, so that the payoff to each resident of the grand coalition is

$$(1 + \varepsilon)ea - \frac{1}{2}be^2(1 + 2\varepsilon) \quad (29)$$

In a two-bloc coalition structure, in which one large bloc has a share $1 - s_N$ of the population and the other, smaller bloc has a share s_N , the payoff to the smaller bloc must be a decreasing function of s_N , as Lemma 10 below shows. If $s_N = 0.5$, the payoff to the small bloc in this deviating coalition must be less than expression (29) ; if s_N were so small that condition A2 held with equality, then this payoff would exceed (29) ; therefore there is some threshold $s^0 < 0.5$, such that the payoff to the smaller bloc in this deviating 2-bloc coalition structure equals (29) (and exceeds it whenever $s_N > s^0$).

Therefore, to prove the theorem it remains to show that no resident of any jurisdiction, in any coalition structure, can do better than a resident of jurisdiction N does in the deviating coalition structure.

First, attention can be restricted to structures with 2 coalitions. If a structure had more than 2 coalitions, Proposition 4 shows that the smallest coalition gets the highest payoff in any structure. Proposition 5 shows that this coalition would do better yet if all of the other coalitions in the structure merged.

Thus it remains to prove that jurisdiction 2, in the deviating coalition structure, does at least as well as the smaller coalition in any other 2-coalition structure. Let (ζ, t_1, t_2) denote the index of tax competition, and the tax rates, in the deviating coalition structure, and let (ζ', t'_1, t'_2) denote the analogous values in any other 2-coalition structure. Proposition 2 implies that $\zeta \geq \zeta'$. Since the share of population in the larger coalition cannot be larger in the other structure than in the deviating structure, equation (14) shows that $t_1 \geq t'_1$. Lemma 3 shows that coalition 2 gains from a tax increase by coalition 1. Therefore, if jurisdiction n were to set its tax rate at t'_2 , when the larger coalition in the deviating structure sets a tax rate of t_1 , jurisdiction N must do at least as well as the smaller coalition does in the other coalition structure. And its payoff when it chooses its best response t_2 to t_1 must be at least as high as its payoff when it chose t'_2 and the other coalition in the deviating structure chose t_1 . Therefore, the payoff to jurisdiction N in the deviating coalition structure is at least as high as the payoff to any coalition in any structure (other than the coalition of the whole), completing the proof of the proposition. •

Lemma 8 *Under Assumption A2, the return to capital is positive in the deviating coalition structure.*

Proof Proposition 1 shows that ζ is an increasing function of ε , so it is sufficient to prove that condition A2 implies a positive return to capital when $\varepsilon \rightarrow \infty$. When $\varepsilon \rightarrow \infty$,

$$\zeta = \frac{\alpha}{1 - \alpha}$$

so it is sufficient to prove that $a(1 - \alpha) > be$ under assumption A2 where With $\varepsilon = \infty$,

$$\alpha = \frac{s_N}{2 - s_N} + \frac{1 - s_N}{1 + s_N}$$

so that

$$1 - \alpha = \frac{3s_N(1 - s_N)}{(1 + s_N)(2 - s_N)}$$

and assumption A2 implies that $a(1 - \alpha) > be$. •

Lemma 9 *Under assumption A1, the grand coalition wishes to set $t = a - be$.*

Proof Suppose the federation set its tax rate t greater than $a - be$, so that some capital was withheld. Since the net return to capital must be zero in such a case,

$$t = a - bk$$

Therefore, with capital withheld, so that $k < e$, the federation has

$$g = (a - bk)k$$

and

$$x = w = \frac{1}{2}bk^2$$

implying a payoff to each resident of

$$u = \frac{1}{2}bk^2 + (1 + \varepsilon)k(a - bk) = (1 + \varepsilon)ak - bk^2\left(\frac{1}{2} + \varepsilon\right)$$

so that

$$\frac{\partial u}{\partial k} = a(1 + \varepsilon) - be(1 + 2\varepsilon)$$

at $k = e$. Assumption A1 then implies that this derivative is positive at $k = e$ (and at $k < e$), so that the federation would not want to set a tax rate so high that capital was withheld. •

Lemma 10 *In a 2-jurisdiction world, the payoff to the smaller jurisdiction is a decreasing function of its share of the world population.*

Proof Let $s \leq 0.5$ be the population share of the smaller jurisdiction, so that $1 - s$ is the share in the larger jurisdiction. Proposition 2 shows that the index ζ of tax competition is a decreasing function of s . Equation (14) shows that the tax rate t_1 in the larger of the two jurisdictions is an increasing function of its own population share $1 - s$, as well as of ζ . Therefore, as s falls, t_1 must rise.

Lemma 3 shows that the payoff π_2 to the smaller jurisdiction is an increasing function of the tax rate t_1 in the larger jurisdiction. Since jurisdiction 2's own tax rate is its best response to t_1 , it must be the case that jurisdiction 2's equilibrium payoff must increase as its own share of population decreases. •

References

- [1] F. Bloch. Sequential formation of coalitions in games with externalities and fixed payoff division. *Games and Economic Behavior*, 14:90–123, 1996.
- [2] M. Brühlhart and M. Jametti. Vertical versus horizontal tax externalities : An empirical test. *Journal of Public Economics*, 90:2027–2062, 2006.
- [3] S. Bucovetsky. Asymmetric tax competition. *Journal of Urban Economics*, 30:167–181, 1991.
- [4] J. Burbidge, J. de Pater, G. Myers, and A. Sengupta. A coalition–formation approach to equilibrium federations and trading blocs. *American Economic Review*, 87:940–956, 1997.
- [5] P. Dasgupta and E. Maskin. The existence of equilibria in discontinuous economic games I : Theory. *Review of Economic Studies*, 53:1–26, 1986.
- [6] M. Devereux, B. Lockwood, and M. Redoano. Do countries compete over corporate tax rates? *Journal of Public Economics*, forthcoming.
- [7] L. Grazzini and T. van Ypersele. Fiscal coordination and political competition. *Journal of Public Economic Theory*, 5(2):305 – 325, 2003.
- [8] N. Hansen and A. Kessler. The political geography of tax h(e)avens and tax hells. *American Economic Review*, 91:1103–1115, 2001.
- [9] S. Hart and M. Kurz. Endogenous formation of coalitions. *Econometrica*, 51:1047–1064, 1983.
- [10] H. Hwang and B. Choe. Distribution of factor endowments and tax competition. *Regional Science and Urban Economics*, 25(5):655–673, 1995.
- [11] R. Kanbur and M. Keen. Jeux sans frontieres : Tax competition and tax coordination when countries differ in size. *American Economic Review*, 83:877–892, 1993.
- [12] M. Keen and S. Lahiri. The comparison between destination and origin principles under imperfect competition. *Journal of International Economics*, 45, 323 – 350.

- [13] D. Laussel and M. Le Breton. Existence of Nash equilibria in fiscal competition models. *Regional Science and Urban Economics*, 28:283–296, 1998.
- [14] R. Musgrave. Devolution, grants, and fiscal competition. *Journal of Economic Perspectives*, 11(4):65 – 72, Fall 1997.
- [15] W. Oates. *Fiscal Federalism*. Harcourt Brace Jovanovich, New York, 1972.
- [16] S. Peralta and T. van Ypersele. Coordination of capital taxation among asymmetric countries. Technical Report 3695, CEPR, 2003.
- [17] S. Peralta and T. van Ypersele. Factor endowments and welfare levels in an asymmetric tax competition game. *Journal of Urban Economics*, 57(2):258 – 274, March 2005.
- [18] S. Peralta and T. van Ypersele. Coordination of capital taxation among asymmetric countries. *Regional Science and Urban Economics*, 36(6):708 – 726, November 2006.
- [19] M. Perry and R. Porter. Oligopoly and the incentive for horizontal merger. *American Economic Review*, 75:219–227, 1985.
- [20] D. Ray and R. Vohra. Equilibrium binding agreements. *Journal of Economic Theory*, 73:30–78, 1997.
- [21] S. Salant, S. Switzer, and R. Reynolds. Losses from horizontal merger : The effects of an exogenous change in industry structure on Cournot–Nash equilibrium. *Quarterly Journal of Economics*, 98:185–199, 1983.
- [22] M. Sato. Tax competition, rent seeking, and fiscal decentralization. *European Economic Review*, 47(1):19 – 40, February 2003.
- [23] S. Stöwhase. Asymmetric capital tax competition with profit shifting. *Journal of Economics*, 85(2):175 – 196, August 2005.
- [24] D. Wildasin. Nash equilibrium in models of tax competition. *Journal of Public Economics*, 35:229–240, 1988.

- [25] J. Wilson. A theory of interregional tax competition. *Journal of Urban Economics*, 19:296–315, 1986.
- [26] J. Wilson. Tax competition with interregional differences in factor endowments. *Regional Science and Urban Economics*, 21:423–451, 1991.
- [27] J. Wilson. Theories of tax competition. *National Tax Journal*, 52:269–304, 1999.
- [28] G Zodrow and P. Mieszkowski. Pigou, Tiebout, property taxation and the underprovision of local public goods. *Journal of Urban Economics*, 19:356–370, 1986.