IMPLEMENTING TAX COORDINATION
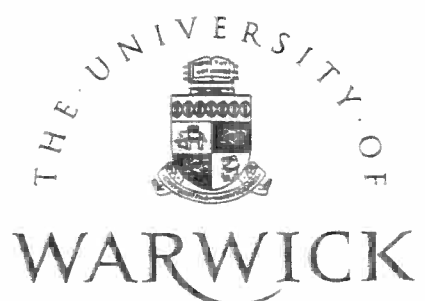
Amrita Dhillon, Carlo Perroni and Kimberley A. Scharf

No.501

WARWICK ECONOMIC RESEARCH PAPERS

THE UNIVERSITY OF WARWICK

DEPARTMENT OF ECONOMICS

IMPLEMENTING TAX COORDINATION

Amrita Dhillon
Department of Economics
University of Warwick
Coventry, CV4 7AL
UK

Carlo Perroni
Department of Economics
University of Warwick
Coventry, CV4 7AL
UK

Kimberley A. Scharf
Department of Economics
University of Warwick, Coventry
&
Institute of Fiscal Studies, London

No.501

This version: November 1997

Implementing Tax Coordination[*]

Amrita Dhillon
University of Warwick

Carlo Perroni
University of Warwick

Kimberley A. Scharf
University of Warwick and Institute for Fiscal Studies

This version: November 1997

Abstract

This paper investigates whether tax competition can survive under tax coordination, when information is private or nonverifiable. We focus on a two-jurisdiction model where capital can move across borders, and where jurisdictions have different public good requirements, but are otherwise identical. In this setting, coordination may call for a second-best allocation supported by differentiated tax rates. If, however, information on jurisdictions' types is private or nonverifiable, such a second-best allocation may not be implementable. We show that incentive compatibility requirements will generally affect not only the choice of coordinated rates in states where jurisdictions are different, but also the choice of harmonized rates in states where jurisdictions have identical preferences for public consumption.

KEY WORDS: Tax Competition, Tax Coordination, Implementation.

JEL CLASSIFICATION: H2, H7.

Correspondence should be addressed to Kim Scharf, Department of Economics, University of Warwick, Coventry CV4 7AL, U.K. E-mail: K.Scharf@warwick.ac.uk.

---

1 Introduction

The theoretical literature on tax competition has pointed out that, in a multijurisdictional world, independent fiscal choices can lead to policy coordination failure. In the case of capital income taxation, this failure takes the form of suboptimal levels of taxation and public good provision (Gordon and Wilson, 1986; Zodrow and Mieszkowski, 1986); with respect to commodity taxation, the noncooperative outcome is characterized by equilibrium rate structures that depart from optimal tax rules (Mintz and Tulkens, 1986; Lockwood, 1993). The overall conclusion from this literature is that there could be gains from coordination.

Even if one abstracts from the problem of how cooperation among sovereign jurisdictions can be sustained, there remains the question of which form such coordination should take. With respect to commodity taxation, several studies have examined whether there exist Pareto-improving harmonizing reforms (Keen, 1987, 1989; Turunen-Red and Woodland, 1990; de Crombrugghe and Tulkens, 1990; Lahiri and Raimondos-Møller, 1996). The notion of harmonization (i.e., of a move towards a common system of taxes), however, is rather more restrictive than coordination. If jurisdictions are sufficiently different—with respect to their preferences, technologies or other characteristics—then a second-best allocation may call for differentiated taxes across jurisdictions. In this case, tax harmonization, in its narrower definition, may not be desirable (Kanbur and Keen, 1993); what may be required instead is a form of coordination that takes the different needs of different jurisdictions into account.

The downside of this broader approach to coordination is that it may create incentives for individual jurisdictions to misrepresent their needs, thus making coordination ineffective. This problem is well illustrated by the recent requests for special tax treatment that have come to the fore in the European context—the latest example being represented by France's call for special tax incentives for the region of Ile-de-France, following Jacques Chirac's electoral promises. Such requests have been interpreted as falling under the umbrella of Paragraphs 2 and 3 of Article 92 of the EC Treaty granting special waivers to the ban against "State Aid" to certain geographical areas which, for economic or social reasons, are deemed to be in need of targeted incentives. There are fears that such provisions may lead to

1

a proliferation of calls for special treatment by individual regions. As the harmonization of member states' tax systems moves forward, this trend could spread further and ultimately undermine tax coordination efforts by the center.

This possibility of a re-emergence of competitive pressures in a coordinated fiscal environment is the focus of our study. Specifically, we examine the implementation problem associated with tax coordination in an economic union, where capital is interjurisdictionally mobile and where coordination can only be pursued through the establishment of common rules, rather than through delegation of discretionary power. In such a setting, coordinating choices must be conditioned—via a set of agreed rules—on actions taken by the jurisdictions themselves, which the coordinating authority cannot overrule on the basis of direct observation of the jurisdictions' characteristics, even if these are perfectly observable. This type of institutional constraint is analogous to that underlying traditional analyses of optimal income taxation, where it is assumed that taxes must be income based independently of whether or not individuals' abilities are observable; this, in turn, is formally analogous to examining a scenario where information about types is private.[1]

We are concerned with situations where jurisdictions voluntarily enter into a binding coordination arrangement for the sole purpose of overcoming the fiscal coordination failure associated with tax competition, and where this objective is pursued through the delegation of the choice of common rules for selecting tax rates to a coordinating body, which is given no mandate to engage in redistributive policies, nor to independently levy taxes. We focus on an environment where jurisdiction types are fully uncorrelated, and where jurisdictions can simultaneously attempt to "cheat" on the tax coordination scheme. This framework enables us to capture the idea of a simultaneous race for preferential tax treatment by member regions of a coordinated union, and investigate whether tax competition can re-emerge in the coordinated solution as the result of the simultaneous threat of misrepresentation by individual jurisdictions; and if so, which form it will take.

The basic framework of our analysis is a two-jurisdiction model of tax competition where the two jurisdictions have different public good requirements but are otherwise identical. Public good provision in each jurisdiction is financed through a general output tax—which

in our model can be thought of as representing origin-based commodity taxation or, equivalently, source based comprehensive income taxation. Such form of taxation is generally distortionary, unless rates are set equally across jurisdictions, and in the absence of coordination would lead to tax competition and suboptimal levels of taxation. We show that unconstrained coordination of tax polices may call for the choice of a second-best allocation supported by differentiated tax rates for the two jurisdictions; if, however, information on jurisdiction types is nonverifiable, such a second-best allocation may not be implementable. We show that, in these cases, the presence of incentive-related constraints can affect not only the choice of coordinating rates in states where jurisdictions are different, but also the choice of a harmonized rate in states where jurisdictions have identical public good requirements.

The plan of the paper is as follows. Section 2 describes our basic setup, and Section 3 discusses tax coordination under full information. Section 4 examines how tax coordination can be attained under Bayesian implementation. Section 5 summarizes and concludes.

2 Taxation and capital flows

Consider an economy with two jurisdictions having identical endowments and technologies. Each jurisdiction is endowed with a fixed amount of labour, which is interjurisdictionally immobile, and with one unit of capital, which can move freely across jurisdictions. Each jurisdiction $i \in \{1, 2\}$ produces a private good and a public good using labour and capital. Trade in this economy consists of a flow of capital inputs from one jurisdiction to the other and a corresponding flow of produced goods in the opposite direction.

We shall assume the marginal rate of transformation between the private and public good to be constant and equal to unity. We can make the presence of the immobile factor implicit by representing technologies by means of a concave production function f, whose argument is capital inputs k_i ($i \in \{1, 2\}$). For the sake of simplicity, we shall assume the inverse elasticity of demand for capital $\eta \equiv f'[k_i]/(k_i f''[k_i]) < 0$ to be constant. Without loss of generality, we will also assume $f[1] = 1$.

3

There exist two possible types, $\overline{\theta}$ and $\underline{\theta}$, for each of the two jurisdictions. These two types have identical technologies and endowments but differ in their preferences for private and public goods. The preferences of the representative consumer in each jurisdiction are summarized by a quasilinear utility function:

$$u_i[c_i, g_i \mid \theta_i] \equiv c_i + \theta_i h[g_i], \quad i \in \{1, 2\}, \tag{1}$$

where c_i and g_i are respectively quantities of private and public goods, $h' > 0$, $h'' < 0$, and $\theta_i \in \Theta \equiv \{\overline{\theta}, \underline{\theta}\}$, with $\overline{\theta}$, and $\underline{\theta}$ both greater than zero. We shall assume $\overline{\theta} > \underline{\theta}$, i.e., jurisdictions of type $\overline{\theta}$ value public goods more at the margin.

Suppose that an *ad valorem* output tax is levied in each jurisdiction at rates equal to t_i ($i \in \{1, 2\}$). Let us denote by p the gross-of-tax price of output—which, since goods are traded, must be the same across jurisdictions—and by r_i the return to capital inputs in jurisdiction i ($i \in \{1, 2\}$). Then, profit maximization implies

$$r_i = \frac{p f'[k_i]}{1 + t_i}, \quad i \in \{1, 2\}, \tag{2}$$

i.e., the marginal revenue product of capital must equal its input price. Arbitraging of investment opportunities then implies $r_1 = r_2$, i.e.,

$$\frac{f'[k_1]}{1 + t_1} = \frac{f'[k_2]}{1 + t_2}, \tag{3}$$

and market clearing requires

$$k_1 + k_2 = 2. \tag{4}$$

These conditions identify implicit functions $k_i[t_1, t_2]$ ($i \in \{1, 2\}$).

Normalizing p to unity, we can express private consumption as the net-of-tax difference between domestic output and payments to foreign-source capital:

$$c_i[t_1, t_2] \equiv \frac{f[k_i[t_1, t_2]] - (k_i[t_1, t_2] - 1) f'[k_i[t_1, t_2]]}{1 + t_i}, \quad i \in \{1, 2\}; \tag{5}$$

while public good provision is simply equal to domestic tax revenues:

$$g_i[t_1, t_2] \equiv t_i \frac{f[k_i[t_1, t_2]]}{1 + t_i}, \quad i \in \{1, 2\}. \tag{6}$$

4

Note that allocations (c_1, g_1, c_2, g_2) depend only on the tax rate profile (t_1, t_2), and are independent of jurisdictions' types. We can thus define indirect utility functions as $v_i[t_1, t_2 \,|\, \theta_i] \equiv u_i[c_i[t_1, t_2], g_i[t_1, t_2] \,|\, \theta_i]$ $(i \in \{1, 2\})$. Comparative statics effects with respect to changes in tax rates are presented in the Appendix.

In this model taxation affects capital location decisions through the arbitraging condition (3), and is thus generally distortionary unless tax rates are identical across jurisdictions. Furthermore, if tax policies are uncoordinated, each jurisdiction has an incentive to independently lower its tax rate in order to attract capital, leading to suboptimal levels of taxation and public goods provision.[2] Thus, coordination of tax policies across jurisdictions will generally be called for.

3 Tax coordination under full information

Suppose that, under full information, tax rates are chosen by a central authority, whose only mandate is to oversee tax coordination. In our analysis, we shall maintain the assumption that neither interjurisdictional transfers nor central taxes are available to the coordinating authority. Then, if the two regions are identical, coordination requires the adoption of a common rate across jurisdictions; but if the two jurisdictions are of different types, the adoption of a common rate may not be desirable.[3]

Formally, the coordination problem of the central authority can be described as follows. We shall focus on an *uncorrelated* environment, i.e., a situation where the realization of each jurisdiction's type is independent of the other's type, and assume that each type is equally likely. Thus, there are four possible configurations of types across jurisdictions: $(\bar{\theta}, \bar{\theta})$, $(\underline{\theta}, \underline{\theta})$, $(\bar{\theta}, \underline{\theta})$, $(\underline{\theta}, \bar{\theta})$, which are all equally likely. The coordinating authority must then choose a profile of rates (t_1, t_2) for each of these four states—a total of eight rates.

Let the objective of the coordinating authority be represented by an *ex-ante* symmetric social welfare function W having for arguments the expected utilities of the representative consumers of the two jurisdictions. Because of symmetry, in states where jurisdictions are of the same type the coordinating authority will select identical rates for the two jurisdictions,

and, in the two states where jurisdictions are of different types, the chosen profiles will be symmetric. Thus, the combinations of coordinated tax rates corresponding to each state will be $(\bar{t}, \bar{t})$, $(\underline{t}, \underline{t})$. $(\bar{t}', \underline{t}')$, $(\underline{t}', \bar{t}')$, These can be summarized by four rates: $\bar{t}$—the rate which applies to both jurisdictions when they are both of the high valuation type; $\underline{t}$—the rate which applies to both jurisdictions when they are both of the low valuation type; $\bar{t}'$ and $\underline{t}'$—the rates which apply respectively to the high and low valuation jurisdictions when they are of different types.

We shall assume that the representative consumers of both jurisdictions are risk neutral, which means that there is no scope for risk pooling by the coordinating authority. To be specific, we assume that expected utility is a simple linear aggregation across states of the quasilinear utility function defined in (1):

$$E_1[\bar{t}, \underline{t}, \bar{t}', \underline{t}'] \equiv \frac{1}{4}\left(v_1[\bar{t}, \bar{t}, |\bar{\theta}] + v_1[\underline{t}, \underline{t}, |\underline{\theta}] + v_1[\bar{t}', \underline{t}', |\bar{\theta}] + v_1[\underline{t}', \bar{t}', |\underline{\theta}] \right); \tag{7}$$

$$E_2[\bar{t}, \underline{t}, \bar{t}', \underline{t}'] \equiv \frac{1}{4}\left(v_2[\bar{t}, \bar{t}, |\bar{\theta}] + v_2[\underline{t}, \underline{t}, |\underline{\theta}] + v_2[\underline{t}', \bar{t}', |\bar{\theta}] + v_2[\bar{t}', \underline{t}', |\underline{\theta}] \right). \tag{8}$$

An optimal coordinating choice can thus be characterized as the solution to the problem of maximizing, by choice of rates $\bar{t}, \underline{t}, \bar{t}', \underline{t}'$, the objective

$$W\left[E_1[\bar{t}, \underline{t}, \bar{t}', \underline{t}'], E_2[\bar{t}, \underline{t}, \bar{t}', \underline{t}'] \right]. \tag{9}$$

By symmetry, the expected utilities of the two jurisdictions will be the same for any choice of rates $\bar{t}, \underline{t}, \bar{t}', \underline{t}'$. Thus we can rescale the objective and rewrite it as

$$S[\bar{t}, \underline{t}, \bar{t}', \underline{t}'] \equiv v_1[\bar{t}, \bar{t}, |\bar{\theta}] + v_1[\underline{t}, \underline{t}, |\underline{\theta}] + v_1[\bar{t}', \underline{t}', |\bar{\theta}] + v_1[\underline{t}', \bar{t}', |\underline{\theta}]. \tag{10}$$

As (10) makes clear, this problem is equivalent to maximizing the expectation of ex-post social welfare under a utilitarian social welfare function specification.

The first-order conditions for this problem are

$$\frac{\partial v_1[\bar{t}, \bar{t} \,|\, \bar{\theta}]}{\partial \bar{t}} = 0; \tag{11}$$

$$\frac{\partial v_1[\underline{t}, \underline{t} \,|\, \underline{\theta}]}{\partial \underline{t}} = 0; \tag{12}$$

6

$$\frac{\partial v_1[\bar{t}', \underline{t}' \mid \bar{\theta}]}{\partial \bar{t}'} + \frac{\partial v_1[\underline{t}', \bar{t}' \mid \underline{\theta}]}{\partial \bar{t}'} = 0; \tag{13}$$

$$\frac{\partial v_1[\bar{t}', \underline{t}' \mid \bar{\theta}]}{\partial \underline{t}'} + \frac{\partial v_1[\underline{t}', \bar{t}' \mid \underline{\theta}]}{\partial \underline{t}'} = 0. \tag{14}$$

Note that (11) and (12) are independent of each other and of (13) and (14).

In the two states where jurisdictions are of the same type ($\theta_1 = \theta_2$), and tax rates are chosen to be identical across regions ($t_1 = t_2 = t$), interjurisdictional capital flows are zero, implying $k_1 = k_2 = 1$ and $g_1 = g_2 = t/(1 + t)$. The social marginal cost of public funds (the cost of raising one additional dollar through a simultaneous marginal increase in both rates) is equal to unity, and thus the common optimal rates are characterized by equality between the (common) social marginal valuation of the public good and the (common) social marginal cost of public funds, i.e.,

$$\bar{\theta} h'[\bar{t}/(1 + \bar{t})] = 1; \tag{15}$$

$$\underline{\theta} h'[\underline{t}/(1 + \underline{t})] = 1. \tag{16}$$

Totally differentiating (15) with respect to $\underline{\theta}$ we obtain

$$\frac{\partial \bar{t}}{\partial \bar{\theta}} = -\epsilon \bar{t}(1 + \bar{t}) > 0, \tag{17}$$

where $\epsilon \equiv ((1 + \bar{t})/\bar{t}) \, h'[\bar{t}/(1 + \bar{t})]/h''[\bar{t}/(1 + \bar{t})] < 0$ is the inverse elasticity of the marginal valuation of the public good with respect to public good provision. As would be expected, an increase in the marginal valuation of public goods in both jurisdictions results in a higher optimal tax rate, which implies that the optimal choice of rates features $\bar{t} > \underline{t}$.

In contrast, in states where $\theta_1 \neq \theta_2$, an optimal coordinating choice generally requires different rates across regions. The direction of the deviation from uniformity, however, is ambiguous, depending on technological possibilities as well as preferences. To understand how the coordinating choice departs from uniformity, we can start from the case where $\bar{\theta} = \underline{\theta}$ (where $\bar{t} = \underline{t} \equiv t$) and focus on a marginal increase in $\bar{\theta}$. We can then totally differentiate first-order conditions (13) and (14) with respect to $\bar{t}'$, $\underline{t}'$, and $\bar{\theta}$, and evaluate them at $\bar{\theta} = \underline{\theta} \equiv \theta$ (see Appendix). Letting $q = f'[1]t\eta$, we can state the following proposition:

PROPOSITION 1: *At $\overline{\theta} = \underline{\theta}$, if $|q| < 1$, then $\partial \overline{t}'/\partial \overline{\theta} > \partial \underline{t}'/\partial \overline{\theta}$; i.e., starting from a situation where marginal public good valuations are the same in both jurisdictions, an increase in the marginal valuation of public goods in the high valuation region induces the coordinating authority to raise the high valuation region's tax rate relative to the other region's. If $|q| > 1$, then $\partial \overline{t}'/\partial \overline{\theta} < \partial \underline{t}'/\partial \overline{\theta}$; i.e., an increase in the marginal valuation of public goods in the high valuation region induces the coordinating authority to lower the higher valuation region's tax rate relative to the other region's.*

See the Appendix for a proof.

Thus, when jurisdictions are of different types, an optimal coordinating choice requires unequal tax rates across the two jurisdictions. This will distort capital allocation decisions, preventing the attainment of productive efficiency.[4] Proposition 1 shows that, while in general this involves higher taxes for the jurisdiction with a higher θ, if capital is highly mobile (i.e., when q, which is proportional to η, is large in absolute value) we may see higher taxes for the jurisdiction with a lower marginal valuation. This latter "perverse" result can be explained as follows. On efficiency grounds, an increase in $\overline{\theta}$ generates pressure to raise public good provision in the high valuation region. When η is sufficiently large in absolute value (i.e., capital is highly mobile) it may be possible to achieve this result by raising the tax rate in low valuation region relative to the other region's; this causes capital to move to the high valuation region, bringing about an enlargement of the tax base in the high valuation region; such enlargement could be sufficient to compensate for the lower tax rate, and result in increased tax revenues in the high valuation region.

4 The implementation problem

Now suppose that the coordinating authority cannot independently choose taxes on the basis of observation of the jurisdictions' requirements. Instead, it must establish a set of rules that condition tax rates on actions taken by the jurisdictions themselves. Then, the coordinating authority faces an implementation problem, i.e., it cannot directly choose a

given outcome, and must devise a mechanism such that the two jurisdictions, by their choice of strategies, reach the desired outcome.

As discussed earlier, this problem is formally identical to implementing optimal choice rules under private information: hence, in the remainder of our analysis we shall adopt terminology and ideas which apply to an asymmetric information scenario. What is distinctive about the problem at hand is that, since the only available policy instruments are output taxes, it is not possible to devise a mechanism which directly maps strategies into allocations; rather, allocations must be supported by an appropriate choice of tax rates. And since output taxes are distortionary in this model, it will only be possible to attain a subset of all the allocations that are compatible with the resource constraint. Thus, the second-best nature of the instrument available to the coordinating authority makes this a "constrained" implementation problem.

Calling upon the revelation principle (Dasgupta, Hammond, and Maskin, 1979), we shall restrict our attention to direct mechanisms, whereby the two jurisdictions are given a menu of tax vectors conditioned on the (joint) announcement of types, and such that both jurisdictions are induced to truthfully reveal their type. Formally, the two jurisdictions send messages (μ_1, μ_2) to the central authority. The message space M_i $(i \in \{1, 2\})$ for each jurisdiction equals the set of possible types $(M_i \equiv \Theta, \ i \in \{1, 2\})$, i.e., messages consist of announcements of types. A direct mechanism must specify a mapping from announcement profiles (μ_1, μ_2) to tax profiles (t_1, t_2). In any truth telling equilibrium, we require that each jurisdiction truthfully announces its type when the other jurisdiction does the same. An *optimal* mechanism is one that maximizes expected social welfare.

We shall maintain our previous assumption that the realization of each jurisdiction's type is independent of the other's type. Thus, there are four possible configurations of announcements, $(\overline{\theta}, \overline{\theta})$, $(\underline{\theta}, \underline{\theta})$, $(\overline{\theta}, \underline{\theta})$, $(\underline{\theta}, \overline{\theta})$, each corresponding to a possible configuration of types. As in the previous section, we shall restrict our attention to symmetric mechanisms, where the combinations of tax rates are symmetric across jurisdictions.[5]

We shall also assume that jurisdictions do not observe each other's types (*incomplete information*), and view the two possible type realizations as equally likely. The relevant im-

plementation concept for such an environment is *Bayesian* implementation, which consists of specifying a mapping from announcement pairs to tax rate pairs such that truth-telling is a Bayesian Nash equilibrium. This in turn requires that truth-telling be a best-response "on average" given each jurisdiction's probability assessment of the other's type.[6] Focusing on jurisdiction 1, this requires that the choice of tax rates must satisfy the following two constraints:

$$\overline{\Phi} \equiv \frac{1}{2}\Omega_I + \frac{1}{2}\Omega_{II} \geq 0; \tag{18}$$

$$\underline{\Phi} \equiv \frac{1}{2}\Omega_{III} + \frac{1}{2}\Omega_{IV} \geq 0; \tag{19}$$

where

$$\Omega_I \equiv v_1[\bar{t}, \bar{t} \,|\, \overline{\theta}] - v_1[\underline{t}', \bar{t}' \,|\, \overline{\theta}]; \tag{20}$$

$$\Omega_{II} \equiv v_1[\bar{t}', \underline{t}' \,|\, \overline{\theta}] - v_1[\underline{t}, \underline{t} \,|\, \overline{\theta}]; \tag{21}$$

$$\Omega_{III} \equiv v_1[\underline{t}, \underline{t} \,|\, \underline{\theta}] - v_1[\bar{t}', \underline{t}' \,|\, \underline{\theta}]; \tag{22}$$

$$\Omega_{IV} \equiv v_1[\underline{t}', \bar{t}' \,|\, \underline{\theta}] - v_1[\bar{t}, \bar{t} \,|\, \underline{\theta}]. \tag{23}$$

The expressions Ω_I and Ω_{II} represent the payoff difference for jurisdiction 1 between announcing truthfully and lying, if it is of the high type, and given that jurisdiction 2 announces truthfully, and is respectively of the high and of the low type. The expressions Ω_{III} and Ω_{IV} represent the payoff difference for jurisdiction 1 between announcing truthfully and lying, if it is of the low type, and given that jurisdiction 2 announces truthfully, and is respectively of the high and of the low type. The first constraint (18) thus says that the expected payoff of jurisdiction 1 from announcing its true type must be higher than the corresponding expected payoff from lying, if it is of the high valuation type, while the second constraint (19) says that the expected payoff of jurisdiction 1 from announcing its true type must be higher than the corresponding expected payoff from lying, if it is of the low valuation type.

Since the incentive compatibility constraints for jurisdiction 2 are symmetrically identical to those for jurisdiction 1, and since they impose the same restrictions on the choice

of rates, we can neglect them and focus on the above two constraints. We can also neglect the participation constraints: since, by construction, the coordinated solution is at least weakly Pareto superior (in terms of expected utility) to a symmetric noncooperative outcome, jurisdictions will always be better off by participating in the scheme.[7]

4.1 Implementing the full information choice: differential analysis

Can the full information optimal tax structure be implemented in the presence of these additional constraints?[8] We will begin addressing this question by again examining a small perturbation of types, starting from a situation where the two types are identical (i.e., $\overline{\theta} = \underline{\theta}$) and moving to a situation where the two types are only slightly different (the discrete case will be examined later). Note that when $\overline{\theta} = \underline{\theta}$ (and provided this is common knowledge), the incentive compatibility constraints become irrelevant. Full information optimal tax rates are all identical and equal to t; for $\overline{\theta} = \underline{\theta}$, such a configuration of rates also satisfies all incentive compatibility constraints with equality. Thus the full information optimal choice of rates can always be implemented.

Now consider a marginal increase in $\overline{\theta}$ under full information. The direction of full information tax reform will be characterized by a vector $(\partial \overline{t}'/\partial \overline{\theta}, \partial \underline{t}'/\partial \overline{\theta}, \partial \overline{t}/\partial \overline{\theta})$, whose elements correspond to the expressions we derived in Section 3 ((17) and (51)-(52) in the Appendix). We can then ask the following question: if information is private or nonverifiable, will the full information direction of reform be compatible with truthful revelation? If we evaluate $\overline{\Phi}$ and $\underline{\Phi}$ when types are identical at the full-information optimal choice of rates, both are (weakly) non-binding. As we perturb $\overline{\theta}$ and move taxes in the direction of full-information tax reform, however, the incentive compatibility constraints will be affected. It is easy to show that such reform violates the incentive compatibility constraints.

PROPOSITION 2: *Starting from $\overline{\theta} = \underline{\theta}$, and following a small perturbation of $\overline{\theta}$, if $|q| \neq 1$ the choice of a full information optimal tax reform vector cannot be implemented in Bayesian Nash equilibrium.*

11

See the Appendix for a proof.

Thus, for a small perturbation of types, the choice of a full information optimal rate structure is not incentive compatible. In the "perverse" case where the coordinating authority raises the low valuation type's rate above the high valuation type's rate, the incentive compatibility constraints will be violated for the low valuation type. If, however, the coordinating authority wishes to raise the high valuation type's rate above the low valuation type's rate (the normal case), the high valuation type will misrepresent its type so as to lower its tax rate and attract capital within its boundaries. Thus, in the presence of private or nonverifiable information, capital tax competition reappears in the tax coordination problem in the form of an incentive related constraint on the coordinating choice.

4.2 Optimal mechanisms

As we have just shown, when jurisdictions can be of different types, it will generally not be possible to implement a full information coordinating choice. In such a scenario, the coordinating authority's problem consists of choosing an incentive compatibility constrained optimal mechanism, i.e., choosing rates $\bar{t}, \underline{t}, \bar{t}', \underline{t}'$, which maximize (10) subject to (18) and (19).

We can immediately derive implications for the structure of optimal mechanisms in the neighbourhood of identical types by referring to our previous differential analysis:

PROPOSITION 3: *Starting from $\bar{\theta} = \underline{\theta}$, and following a small perturbation of $\bar{\theta}$, an optimal incentive compatible reform vector will feature $\partial \bar{t}'/\partial \bar{\theta} = \partial \underline{t}'/\partial \bar{\theta}$.*

PROOF: This follows directly from Proposition 2.

Following a local perturbation of types starting from a symmetric scenario, the only incentive compatible reform vectors will be those that maintain $\bar{t}'$ and $\underline{t}'$ at a common level.[9] On the other hand, since local changes in $\bar{t}$ and $\underline{t}$ (the common rates associated

with states where jurisdictions are identical) have no bearing on incentives,[10] the choice of common rates for states where jurisdictions are of the same type will be unaffected. As the following analysis will show, however, when discrete type changes are involved, the choice of common rates can also be affected.

In the following discussion, we will focus on the "normal" case where $\bar{t}' > \underline{t}'$. Four different regimes could conceivably arise at such an optimum: (i) both constraints binding ($\overline{\Phi} = \underline{\Phi} = 0$); (ii) only the first constraint binding ($\overline{\Phi} = 0$, $\underline{\Phi} > 0$); (iii) only the second constraint binding ($\overline{\Phi} > 0$, $\underline{\Phi} = 0$); (iv) neither constraint binding ($\overline{\Phi} > 0$, $\underline{\Phi} > 0$). Our next result states that regime (i) can never occur at an optimum with $\bar{t} > \underline{t}$ and $\bar{t}' > \underline{t}'$.

PROPOSITION 4: *At a constrained optimum with $\bar{\theta} > \underline{\theta}$, $\bar{t} > \underline{t}$, $\bar{t}' > \underline{t}'$, and $g_1[\bar{t}', \underline{t}'] > g_1[\underline{t}', \bar{t}']$, both constraints cannot be simultaneously be binding.*

See the Appendix for a proof.

We can thus rule out regime (i) (both constraints binding). Although our differential analysis has shown that, for local type perturbations, a higher tax rate in the higher valuation jurisdiction ($\bar{t}' > \underline{t}'$) is associated with regime (ii), without more specific assumptions on functional forms we cannot exclude the possibility of regime (iii) occurring. In the remainder of our discussion we shall focus on the second regime (first constraint binding), i.e., on a situation where the high valuation region has an incentive to lie about its true type—which is the case more relevant to the question we are addressing here. In the following section we shall also examine the possibility of regime (iv) (unconstrained optimum) occurring.

The first constraint consist of two components, represented by expressions Ω_I and Ω_{II} defined in (20) and (21). Our previous differential analysis has shown that these two expressions are locally equivalent (see the proof of Proposition 2); but as we move further away from a scenario with identical types this will no longer be the case. The nature of the incentive problem associated with each half of the constraint can be illustrated with the help of Figures 1 and 2. On the horizontal axis we depict the first jurisdiction's rate and

13

on the vertical axis the second jurisdiction's rate. We can then draw indifference curves for the indirect utility functions $v_1[t_1, t_2 \mid \theta_1]$ and $v_2[t_1, t_2 \mid \theta_2]$.

Figure 1 illustrates the incentive problem associated with the first half of the constraint (represented by the expression Ω_I in (20)). Consider a situation where both jurisdictions are of the high valuation type. The full information optimal choice of rates will be represented by a point on the 45 degree line where the first-order conditions for a social optimum (13)-(14) are satisfied. These conditions can be rearranged to give

$$-\frac{\partial v_1[t_1, t_2 \mid \theta_1]/\partial t_1}{\partial v_1[t_1, t_2 \mid \theta_1]/\partial t_2} = -\frac{\partial v_2[t_1, t_2 \mid \theta_2]/\partial t_1}{\partial v_2[t_1, t_2 \mid \theta_2]/\partial t_2}. \qquad (24)$$

The expressions on either side of (24) represent the slopes of the indirect utility indifference curves; thus, condition (24) simply states that an optimum is characterized by tangency between two indifference curves. In the Appendix (proof of Proposition 1) we show that, at a symmetric optimum, $\partial v_1[t_1, t_2 \mid \theta_1]/\partial t_1 = -\partial v_1[t_1, t_2 \mid \theta_1]/\partial t_2 = q/(2(1+t)^2)$, where t is the common rate; both expressions in (24) are therefore equal to 1, implying that the common slope of the two indifference curves is equal to unity at an optimum. Furthermore, the second-order conditions for an optimum require that the indifference curve for jurisdiction 1 be convex to the 45 degree line from its left, while the indifference curve for jurisdiction 2 must be convex to the same line from its right.[12] We also know that the indirect utility of each region is monotonically increasing in the other region's tax rate, which implies that points in the shaded area of Figure 1 are preferred by jurisdiction 1 to point A.

Suppose that θ_1 falls while θ_2 remains unchanged, and suppose that the reaction of the coordinating authority is to lower t_1 relative to t_2 (the normal case). This choice will be represented by a point to the left of the 45 degree line such as B, which also corresponds to a tangency point between an indifference curve of jurisdiction 2 (when of the high valuation type) and an indifference curve of jurisdiction 1 (when of the low valuation type). If, however, this point lies in the area above the high valuation indifference curve for jurisdiction 1 (going through A), the first jurisdiction, if it is of the high valuation type, would have an incentive to misrepresent its type to reach point B. The coordinating authority would then be forced to choose a point such as B', which makes jurisdiction 1 indifferent between lying and telling the truth. Note that the coordinating authority will never attempt to
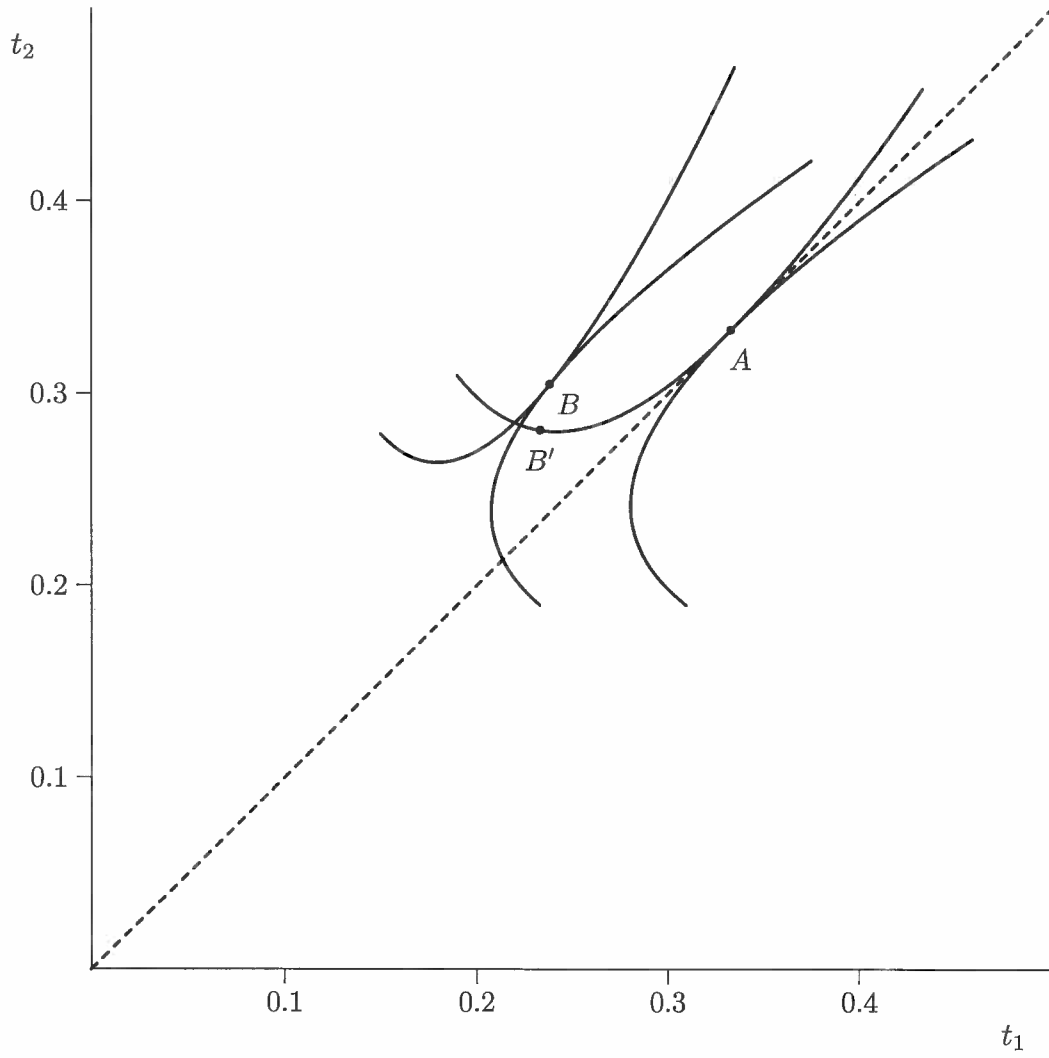
Figure 1: Incentives when $\theta_1 = \theta_2 = \overline{\theta}$.[11]

manipulate incentives by changing its choice of common rate (point A); this is because any other point along the 45 degree line makes jurisdiction 1 worse off (we shall prove this formally below).[13]

The incentive problem associated with the second half of the constraint (the expression Ω_{II} in (21)) is illustrated in Figure 2. Consider now a situation where jurisdiction 2 is of the high valuation type and 1 of the low valuation type, and such that the full information coordinating choice involves a higher tax rate for jurisdiction 2—a point such as B to the left of the 45 degree line. Now compare this with a situation where both jurisdictions are of the low valuation type, which under full information would lead to the choice of a common tax rate, represented by a point on the 45 degree line such as C. If this point lies in the area below the high valuation indifference curve for jurisdiction 2 (going through B), then jurisdiction 2, if it is of the high valuation type, would have an incentive to announce it is of the low valuation type. The coordinating authority would then be forced to choose a point such as C'.

The implementation problem we are analyzing simultaneously involves both types of incentive problems for both jurisdictions, effectively linking the choice of points B' and C': the choice of point B' affects the second half of incentive compatibility constraint (Figure 2), making it less attractive for a high valuation jurisdiction to lie when the other jurisdiction is of the low valuation type; which could cause the second half of the constraint to be positive at an optimum. Nevertheless, it can be shown that, at a constrained optimum, the choice of C' will always be affected:

PROPOSITION 5: *In a constrained optimum with $\overline{\Phi} = 0$ under Bayesian implementation, the choice of $\overline{t}$—the common rate for states where both jurisdictions are of the high valuation type—coincides with the full information choice; $\underline{t}$—the common rate for states where both jurisdictions are of the low valuation type—is less than the full information optimal choice.*
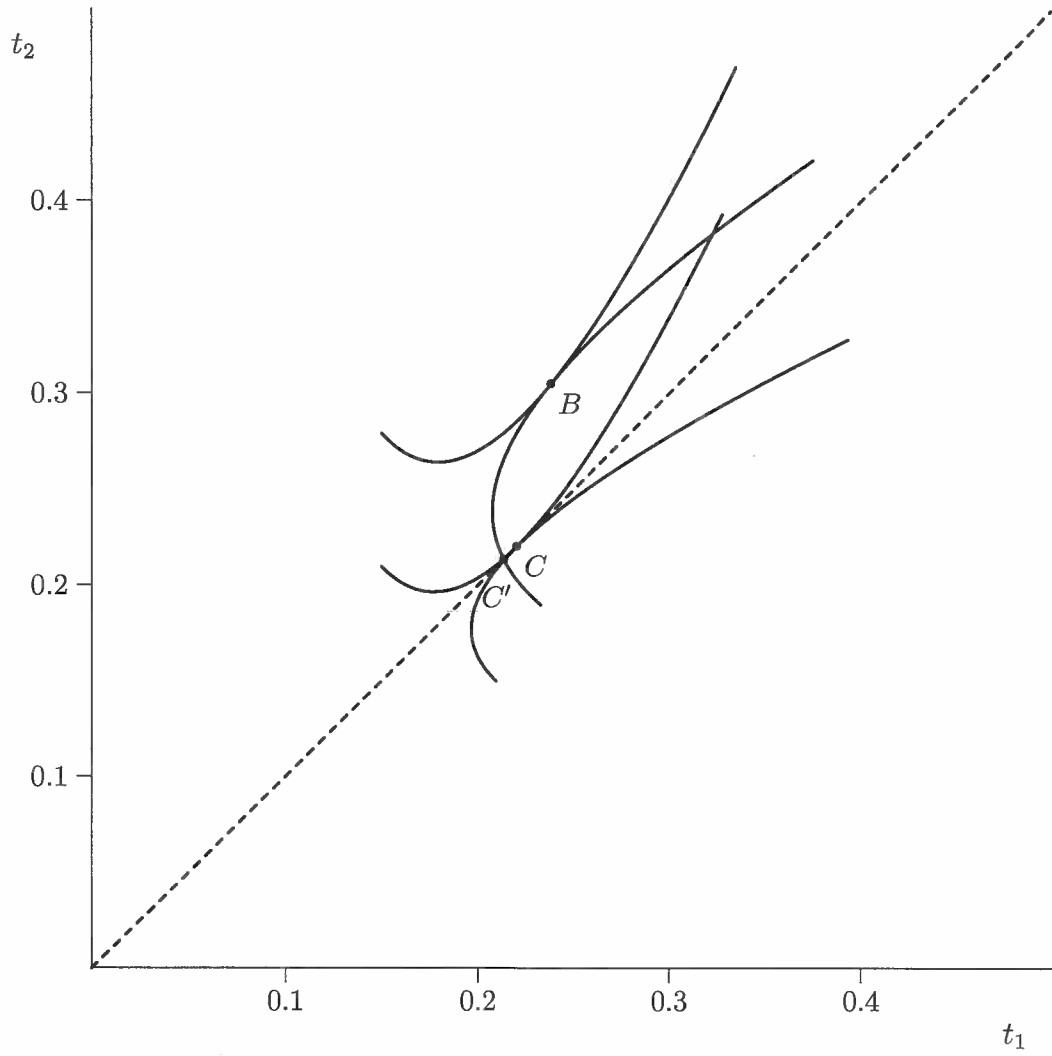

See the Appendix for a proof.

Figure 2: Incentives when $\theta_1 = \underline{\theta}$ and $\theta_2 = \overline{\theta}$.[14]

17

The intuition for why the choice of $\bar{t}$ is unaffected has already been discussed with reference to Figure 1. The intuition for why the choice of $\underline{t}$ is affected is as follows. Under Bayesian implementation, the high valuation type's incentives for truth-telling are based on its *expected* return from telling the truth relative to lying. Suppose that the high valuation jurisdiction gains from lying if the other jurisdiction is of the high type, but loses if the other jurisdiction is of the low type. The loss experienced in the second state discourages a high valuation jurisdiction from lying; by lowering $\underline{t}$, the coordinating authority can make such loss more severe and induce truth-telling across states. In other words, by lowering this common rate the coordinating authority can "punish" simultaneous attempts to lie by both jurisdictions; which, in turn, enables it to better differentiate across jurisdictions in states where they are of different types.

Thus, under Bayesian implementation, incentive compatibility requirements can affect not only the choice of rates for states where the types are different, but also the choice of a common rate for states where both jurisdictions are of the low valuation type: not only can incentive requirements induce the coordinating authority to differentiate less across jurisdictions having different needs, but they can also bring about a "race to the bottom" in the coordinated outcome.

We can illustrate the implications of incentive compatibility requirements for a constrained optimum coordinating choice with the help of an example. Let $f[k_i] = (k_i)^{1/2}$ (which implies $\eta = -2$), $h[g_i] = (g_i)^{1/2}$ (which implies $\epsilon = -2$), $\bar{\theta} = 1$, $\underline{\theta} = 0.85$. The full information solution is $\bar{t} = 33.3\%$, $\bar{t}' = 30.5\%$, $\underline{t}' = 23.8\%$, $\underline{t} = 22\%$.[15] The incentive compatible optimal choice of differentiated rates is $\bar{t}' = 29.4\%$, $\underline{t}' = 24.3\%$, which are lower on average and closer to each other than the corresponding full information rates. The optimal choice of $\bar{t}$ is still 33.3% as in a full information scenario, whereas the constrained optimal choice of $\underline{t}$ is 21.4%; as Proposition 5 indicates, this is less than full information choice.

4.3 Implementing the full information choice: discrete type changes

Can we rule out regime (iv), where neither constraint is binding? It turns out that we cannot. We know that in the neighbourhood of an optimum the indifference curves of both jurisdictions are positively sloped (Figure 3). We also know that an increase in jurisdiction 2's tax rate will raise jurisdiction 1's utility; and that, starting from an optimum, a small decrease in jurisdiction 1's own rate will also raise its welfare. But, for a given t_2, below a certain level of t_1 a further decrease actually lowers jurisdiction 1's utility. Therefore the indifference curves of jurisdiction 1 bend backwards.[17] In turn, this implies that, if the jurisdictions' types are sufficiently different, when jurisdiction 1 is of the high valuation type, point B will lie on a lower indifference curve than point A, and any incentive for jurisdiction 1 to misrepresent its type will vanish (see Figure 3).

PROPOSITION 6: *If $\bar{\theta}$ is sufficiently larger than $\underline{\theta}$, then it is possible to implement the full information choice of rates in Bayesian Nash equilibrium.*

PROOF: We can prove Proposition 6 by means of an example. Consider the case discussed at the end of Section 4.2. There we showed that with $\underline{\theta} = 0.85$ a full information choice is not implementable in Bayesian Nash equilibrium. If we lower $\underline{\theta}$ to 0.75, however, it is possible, under Bayesian implementation, to attain the full information choice of $\bar{t} = 33.3\%$, $\bar{t}' = 29.1\%$, $\underline{t}' = 18.3\%$, and $\underline{t} = 16.4\%$.

Thus, if jurisdictions' types are different but close, the full information choice cannot be implemented (Proposition 2); but if jurisdictions' types are sufficiently different, incentive compatibility requirements have no impact on the coordinated solution. This result, however, should not be overemphasized, as it follows from our rather artificial assumption that only one high type and one low type are possible. In practice, a coordinator would likely be facing a multiplicity of possible type realizations in a given range. If the possible type realizations are sufficiently close to each other, a full information choice would not be implementable even if this range is large.
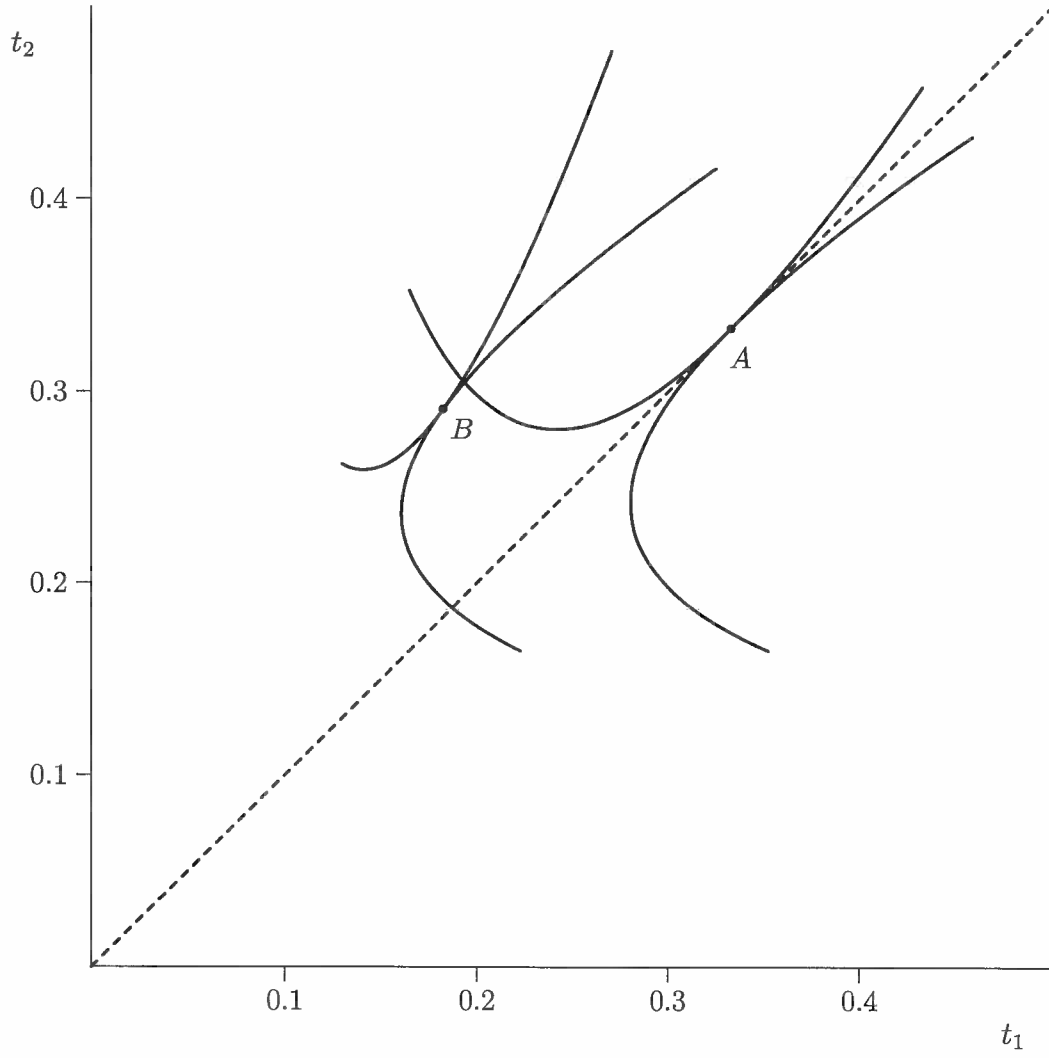
Figure 3: Incentives when $\theta_1 = \theta_2 = \overline{\theta}$, and $\overline{\theta} - \underline{\theta}$ large.[16]

5 Concluding remarks

This paper has been concerned with the possibility that a "race to the bottom" in taxes could take place even when fiscal choices are coordinated. We have investigated the implementation constraints affecting coordination of tax policies in an economic union where jurisdictions have different public good requirements. In this setting coordination may call for a second-best allocation supported by differentiated tax rates for the two jurisdictions, but such a choice will not generally be implementable when information is private or non-verifiable.

Our analysis shows that the presence of incentive-related constraints can affect not only the choice of coordinating rates in states where jurisdictions are different, but also the choice of harmonized rates in states where jurisdictions are identical. Thus, tax competition can survive in disguise even when fiscal policies are coordinated.

These results are particularly relevant to the current debate on tax coordination in Europe, where coordination of economic policies has traditionally been pursued through the definition of rules, rather than through delegation of powers to the center. Recent proposals by the European Commission seem to confirm this approach: the proposed minimum withholding tax on interest payments (May 1989), which seeks to achieve capital tax coordination through the establishment of a floor for taxes on source based capital income taxes;[18] the recently proposed VAT reform (July 1996), which would involve a move to an origin based principle of taxation, allowing countries some discretion in their choice of tax rates within a given pre-determined rate band. Our findings suggest that rules are inadequate to deal with tax competition: granting discretion to the center could be an essential ingredient of an effective fiscal coordination arrangement.

These findings also suggest a possible alternative interpretation of why support for a fiscal federative structure is often more fragile when jurisdictions are potentially diverse. Such aversion to heterogeneity, which is often perceived as stemming from distributional concerns, could instead be due to concerns about the viability of tax coordination: redistributive pressures could in principle be dealt with through bargaining, but bargaining cannot overcome incentive-related constraints if jurisdictions are heterogeneous.

21

Our analysis could be extended in several directions. As mentioned in the previous section, a natural generalization would be to examine a scenario where there exists a range of possible types; this would give rise to a two-dimensional optimal control problem, similar in structure to a continuous-variable optimal income tax problem (Mirrlees, 1971). Furthermore, we have examined a relatively simple form of tax coordination; the problem we address here with respect to a general output tax would also arise with respect to commodity tax coordination, when jurisdictions trade in multiple goods and factors. Finally, we have assumed that the economic union is isolated with respect to the rest of the world; if factor flows to and from the rest of the world were present, the coordinating authority would be facing a partial coordination problem, where coordination is limited to a subset of jurisdictions, and where the coordinated union maintains a strategic stance *vis à vis* non-members.

Notes

[1]There have been several analyses of the role of private information for the design of optimal redistributive policies in a federation (Bordignon, Manasse and Tabellini, 1996; Cornes and Silva, 1996a,b; Raff and Wilson, 1997), but information is a rather neglected topic in the tax competition literature. Bacchetta and Espinosa (1995) have analyzed tax competition when tax collection and enforcement requires information sharing among countries; Gordon and Bovenberg (1996) have put forward informational asymmetries by investors as a possible explanation for the international immobility of capital and for the existence of tax incentives to inward investment.

[2]The uncoordinated choice for each jurisdiction is obtained by maximizing, by choice of its tax rate, the utility of its representative consumer, given the other jurisdiction's rate. The corresponding first-order conditions define a noncooperative equilibrium:

$$\frac{\partial v_1[t_1, t_2 \mid \theta_1]}{\partial t_1} = 0; \tag{25}$$

$$\frac{\partial v_2[t_1, t_2 \mid \theta_2]}{\partial t_2} = 0. \tag{26}$$

Suppose that $\theta_1 = \theta_2 = \theta$. Then, in a noncooperative equilibrium we have $t_1 = t_2 = t^N$, where t^N solves

$$\left(\frac{\partial v_1[t_1, t_2 \mid \theta]}{\partial t_1} \right)_{t_1=t_2=t^N} = 0; \tag{27}$$

while the common welfare maximizing rate is the rate t^* which solves

$$\left(\frac{\partial v_1[t_1, t_2 \mid \theta]}{\partial t_1} + \frac{\partial v_2[t_1, t_2 \mid \theta]}{\partial t_1} \right)_{t_1=t_2=t^*} = 0. \tag{28}$$

It can be shown that at $t_1 = t_2$ the expression $\partial v_2[t_1, t_2 \mid \theta]/\partial t_1$ is positive (see Appendix), implying $t^* > t^N$.

[3]This does not preclude that a partial move towards harmonization, starting from the noncooperative outcome, might result in a Pareto improvement (Keen, 1989).

[4]If lump-sum transfers between spending authorities in the two jurisdictions were available (but not from spending authorities to consumers), it may be possible for the coordinating authority to

attain a first-best allocation by adopting a common rate and transferring tax revenues from the low valuation, low utility jurisdiction to the high valuation, high utility jurisdiction. In expected utility terms, the availability of such transfers could lead to a Pareto improvement, but it would not eliminate the incentive problem discussed in the introduction; indeed it might create incentives for jurisdictions to overstate their public good requirements.

[5]The use of non-anonymous tax configurations (i.e., rates which depend not only on the jurisdictions' types but also on their identity) would likely not be a politically viable option. Because of this restriction, the revelation principle here means that an *anonymous* direct mechanism can do as well as any *anonymous* indirect mechanism.

[6]A more restrictive implementation concept is *dominant strategy* implementation, which requires that both types have no incentive to lie about their true type, no matter what the other jurisdiction's announcement is. Bayesian implementation imposes less stringent incentive compatibility requirements on the choice of the coordinating authority, implying that this will generally be able to achieve a better outcome than under dominant strategy implementation. In particular, in cases where a full information choice cannot be implemented in dominant strategies, Bayesian implementation may make it possible to attain a full information optimal outcome. For a comparative discussion of dominant strategy and Bayesian implementation, see Laffont and Tirole (1993).

[7]The above is an *ex-ante* notion of participation incentives. One may also wish to require that, once jurisdictions learn what their type is, they should have no incentive to deviate from the scheme (*interim* participation constraints). As mentioned earlier, we are concerned with a scenario where jurisdictions enter into a binding coordinating arrangement, i.e., undertake international obligations which they cannot renege upon *ex-post*—at least in the short run; which makes this second type of constraint not relevant for our analysis.

[8]The implementation literature reserves the term implementation to denote situations where a mechanism makes it possible to fully "mimic" a social choice rule under unrestricted domain (i.e., for all possible type configurations). Here we use the term in a more conventional sense, to describe implementation of specific choices for a given realization of types.

[9]Clearly, such a constrained outcome will result in lower expected welfare relative to the full information choice, although, by continuity, it will still Pareto dominate the uncoordinated outcome.

[10]This is a consequence of the symmetry in the conditions characterizing the initial optimum.

[11]The curves in the figure are generated numerically, and refer to the case $f[k_i] = (k_i)^{1/2}$, $h[g_i] =$

24

$(g_i)^{1/2}$, $\overline{\theta} = 1$, $\underline{\theta} = 0.85$.

[12]Negative semidefiniteness of the Jacobian of the problem requires that the indirect utility function $v_i[t_1, t_2 \,|\, \theta_i]$ be concave in (t_1, t_2).

[13]Moving point A off the 45 degree line (e.g., to a point within the shaded area) is also ruled out because it violates anonymity (and besides, it would have an adverse effect affect on the corresponding incentives for the other jurisdiction).

[14]See Footnote 11.

[15]These and the following values were calculated using numerical optimization methods.

[16]This was generated numerically from the specification used to derive Figures 1 and 2, but with $\underline{\theta} = 0.75$.

[17]The derivative $\partial v_1[t_1, t_2 \,|\, \theta_1]/\partial t_1$ can become negative; see the Appendix.

[18]See Huizinga and Nielsen (1997) for an examination of this proposal.

Appendix

Comparative statics

Total differentiation of (3) and (4) gives

$$\frac{\partial k_1}{\partial t_1} = \frac{\Lambda}{1+t_1} < 0, \tag{29}$$

$$\frac{\partial k_1}{\partial t_2} = -\frac{\Lambda}{1+t_2} > 0, \tag{30}$$

where $\Lambda = \eta k_1 k_2 / (k_1 + k_2) < 0$;

$$\frac{\partial c_1}{\partial t_1} = -\frac{(k_1 - 1)f''[k_1]}{1+t_1}\frac{\partial k_1}{\partial t_1} - \frac{c_1}{1+t_1} \gtrless 0; \tag{31}$$

$$\frac{\partial c_1}{\partial t_2} = -\frac{(k_1 - 1)f''[k_1]}{1+t_1}\frac{\partial k_1}{\partial t_2}; \tag{32}$$

the sign of the above agrees with the sign of $(k_1 - 1)$, which is positive if $t_1 < t_2$, and negative if $t_1 > t_2$.

$$\frac{\partial g_1}{\partial t_1} = \frac{t_1 f'[k_1]}{1+t_1}\frac{\partial k_1}{\partial t_1} + \frac{f[k_1]}{(1+t_1)^2} \gtrless 0; \tag{33}$$

$$\frac{\partial g_1}{\partial t_2} = \frac{t_1 f'[k_1]}{1+t_1}\frac{\partial k_1}{\partial t_2} > 0; \tag{34}$$

$$\frac{\partial v_1}{\partial t_1} = \frac{\partial c_1}{\partial t_1} + \theta_1 h'[g_1]\frac{\partial g_1}{\partial t_1} \gtrless 0; \tag{35}$$

$$\frac{\partial v_1}{\partial t_2} = \frac{\partial c_1}{\partial t_2} + \theta_1 h'[g_1]\frac{\partial g_1}{\partial t_2}; \tag{36}$$

the sign of the above is positive if $t_1 < t_2$. Exchanging subscripts, the above expressions also apply to variables pertaining to jurisdiction 2.

Proof of Proposition 1

Noticing that $v_1[\bar{t}', \underline{t}' \mid \bar{\theta}] = v_2[\underline{t}', \bar{t}' \mid \bar{\theta}]$, we can rewrite the first-order conditions (13)-(14) as

$$\frac{\partial v_1[\underline{t}', \bar{t}' \mid \underline{\theta}]}{\partial \bar{t}'} + \frac{\partial v_2[\underline{t}', \bar{t}' \mid \bar{\theta}]}{\partial \bar{t}'} = 0; \tag{37}$$

$$\frac{\partial v_1[\underline{t}', \bar{t}' \mid \underline{\theta}]}{\partial \underline{t}'} + \frac{\partial v_2[\underline{t}', \bar{t}' \mid \underline{\theta}]}{\partial \underline{t}'} = 0. \tag{38}$$

Differentiation of (37)-(38) with respect to $\bar{t}'$, $\underline{t}'$, and $\bar{\theta}$, yields

$$\left(\frac{\partial^2 v_1}{\partial t_1^2} + \frac{\partial^2 v_2}{\partial t_1^2} \right) d\underline{t}' + \left(\frac{\partial^2 v_1}{\partial t_1 \partial t_2} + \frac{\partial^2 v_2}{\partial t_1 \partial t_2} \right) d\bar{t}' = - \left(\frac{\partial^2 v_1}{\partial t_1 \partial \theta_2} + \frac{\partial^2 v_2}{\partial t_1 \partial \theta_2} \right) d\bar{\theta}; \tag{39}$$

$$\left(\frac{\partial^2 v_1}{\partial t_2 \partial t_1} + \frac{\partial^2 v_2}{\partial t_2 \partial t_1} \right) d\underline{t}' + \left(\frac{\partial^2 v_1}{\partial t_2^2} + \frac{\partial^2 v_2}{\partial t_2^2} \right) d\bar{t}' = - \left(\frac{\partial^2 v_1}{\partial t_2 \partial \theta_2} + \frac{\partial^2 v_2}{\partial t_2 \partial \theta_2} \right) d\bar{\theta}. \tag{40}$$

The partial derivatives of the indirect utility functions v_1 and v_2, evaluated at $\theta_1 = \theta_2 = \underline{\theta}$, with $\bar{t}' = \underline{t}' = t$, are

$$\left(\frac{\partial v_1}{\partial t_1} \right)_{\theta_1 = \theta_2} = \left(\frac{\partial v_2}{\partial t_2} \right)_{\theta_1 = \theta_2} = \frac{q}{2(1+t)^2}; \tag{41}$$

$$\left(\frac{\partial v_1}{\partial t_2} \right)_{\theta_1 = \theta_2} = \left(\frac{\partial v_2}{\partial t_1} \right)_{\theta_1 = \theta_2} = -\frac{q}{2(1+t)^2}; \tag{42}$$

$$\left(\frac{\partial^2 v_1}{\partial t_1^2} \right)_{\theta_1 = \theta_2} = \left(\frac{\partial^2 v_2}{\partial t_2^2} \right)_{\theta_1 = \theta_2} = \frac{q^2 + [(3-t)\epsilon + 4]q + 4}{4t(1+t)^3 \epsilon}; \tag{43}$$

$$\left(\frac{\partial^2 v_1}{\partial t_1 \partial t_2} \right)_{\theta_1 = \theta_2} = \left(\frac{\partial^2 v_2}{\partial t_1 \partial t_2} \right)_{\theta_1 = \theta_2} = -\frac{q^2 + [(1+t)\epsilon + 2]q}{4t(1+t)^3 \epsilon}; \tag{44}$$

$$\left(\frac{\partial^2 v_2}{\partial t_1^2} \right)_{\theta_1 = \theta_2} = \left(\frac{\partial^2 v_1}{\partial t_2^2} \right)_{\theta_1 = \theta_2} = \frac{q^2 + (-1 + 3t)\epsilon q}{4t(1+t)^3 \epsilon}; \tag{45}$$

$$\left(\frac{\partial v_1}{\partial \theta_2} \right)_{\theta_1 = \theta_2} = 0; \tag{46}$$

$$\left(\frac{\partial v_2}{\partial \theta_2} \right)_{\theta_1 = \theta_2} = h[t/(1+t)]; \tag{47}$$

$$\left(\frac{\partial^2 v_1}{\partial t_1 \theta_2} \right)_{\theta_1 = \theta_2} = \left(\frac{\partial^2 v_1}{\partial t_2 \theta_2} \right)_{\theta_1 = \theta_2} = 0; \tag{48}$$

$$\left(\frac{\partial^2 v_2}{\partial t_1 \theta_2} \right)_{\theta_1 = \theta_2} = -\frac{q}{2(1+t)^2}; \tag{49}$$

$$\left(\frac{\partial^2 v_2}{\partial t_2 \theta_2} \right)_{\theta_1 = \theta_2} = \frac{q+1}{2(1+t)^2}; \tag{50}$$

Substituting the above into (39) and (40), we obtain

$$\left(\frac{\partial \underline{t}'}{\partial \overline{\theta}}\right)_{\overline{\theta}=\underline{\theta}} = -\frac{t(1+t)\epsilon}{2} \frac{q^2 + [(1+t)\epsilon+1]q}{q^2 + [(1+t)\epsilon+2]q+1}; \tag{51}$$

$$\left(\frac{\partial \overline{t}'}{\partial \overline{\theta}}\right)_{\overline{\theta}=\underline{\theta}} = \left(\frac{\partial \underline{t}'}{\partial \overline{\theta}}\right)_{\overline{\theta}=\underline{\theta}} + \Delta; \tag{52}$$

where

$$\Delta = -t(1+t)\epsilon \frac{q+1}{q^2 + [(1+t)\epsilon+2]q+1}. \tag{53}$$

The denominator in the above expression agrees in sign with the determinant of the Jacobian, which must be positive for concavity. Thus the sign of Δ agrees with the sign of the expression $q+1$: if $|q| < 1$, then $\partial \underline{t}'/\partial \overline{\theta} < \partial \overline{t}'/\partial \overline{\theta}$, and if $|q| > 1$, then $\partial \underline{t}'/\partial \overline{\theta} > \partial \overline{t}'/\partial \overline{\theta}$.

Proof of Proposition 2

We can differentiate each of the constraints with respect to $\overline{\theta}$ and evaluate the total derivative in the direction of the full-information tax reform vector at $\overline{\theta} = \underline{\theta}$, obtaining (after simplification)

$$\frac{\partial \overline{\Phi}}{\partial \overline{\theta}} = -\frac{\partial \underline{\Phi}}{\partial \overline{\theta}} = \frac{q}{2(1+\underline{t})^2} \left(\frac{\partial \overline{t}'}{\partial \overline{\theta}} - \frac{\partial \underline{t}'}{\partial \overline{\theta}}\right). \tag{54}$$

Earlier we have shown that the sign of the difference within brackets depends on the absolute value of q (Proposition 1). If $|q| < 1$ we have $\partial \overline{t}'/\partial \overline{\theta} > \partial \underline{t}'/\partial \overline{\theta}$, and thus $\partial \overline{\Phi}/\partial \overline{\theta} < 0$, implying that the first constraint is violated; analogously if $|q| > 1$, we have $\partial \underline{\Phi}/\partial \overline{\theta} < 0$, implying that the second constraint is violated.

Proof of Proposition 4

If we add up the expressions $\overline{\Phi}$ and $\underline{\Phi}$, we obtain

$$\overline{\Phi} + \underline{\Phi} = (\overline{\theta} - \underline{\theta}) \left(h\left[g_1[\overline{t}, \overline{t}]\right] - h\left[g_1[\underline{t}, \underline{t}]\right] + h\left[g_1[\overline{t}', \overline{t}']\right] - h\left[g_1[\underline{t}', \overline{t}']\right]\right). \tag{55}$$

28

The expression contained in the second set of brackets is positive if $g_1[\bar{t}', \underline{t}'] > g_1[\underline{t}', \bar{t}']$. Thus, as long as we are in a "nonperverse" scenario with no Laffer type effects, we can conclude that the first and second constraint cannot both be simultaneously binding at an optimum.

Proof of Proposition 5

The Lagrangean for the problem can be written as $S + \beta\overline{\Phi}$. Letting $\gamma \equiv \beta/2$, the first-order conditions for a constrained optimum can be written as

$$(1+\gamma)\frac{\partial v_1[\bar{t}, \bar{t} \mid \overline{\theta}]}{\partial \bar{t}} = 0; \tag{56}$$

$$(1+\gamma)\frac{\partial v_1[\bar{t}', \underline{t}' \mid \overline{\theta}]}{\partial \bar{t}'} + \frac{\partial v_1[\underline{t}', \bar{t}' \mid \underline{\theta}]}{\partial \bar{t}'} - \gamma\frac{\partial v_1[\underline{t}', \bar{t}' \mid \overline{\theta}]}{\partial \bar{t}'} = 0; \tag{57}$$

$$(1+\gamma)\frac{\partial v_1[\bar{t}', \underline{t}' \mid \overline{\theta}]}{\partial \underline{t}'} + \frac{\partial v_1[\underline{t}', \bar{t}' \mid \underline{\theta}]}{\partial \underline{t}'} - \gamma\frac{\partial v_1[\underline{t}', \bar{t}' \mid \overline{\theta}]}{\partial \underline{t}'} = 0; \tag{58}$$

$$\frac{\partial v_1[\underline{t}, \underline{t} \mid \underline{\theta}]}{\partial \underline{t}} - \gamma\frac{\partial v_1[\underline{t}, \underline{t} \mid \overline{\theta}]}{\partial \underline{t}} = 0; \tag{59}$$

$$\overline{\Phi} = 0; \tag{60}$$

$$\gamma > 0. \tag{61}$$

The rate $\bar{t}$ only enters (59); since this is unaffected by the inclusion of the constraint, under Bayesian implementation the choice of $\bar{t}$ will be the same as under full information.

Let $\underline{t}^{FI}$ be the full information optimal rate when $\theta_1 = \theta_2 = \underline{\theta}$. We know that a full information optimum requires $\partial v_1[\underline{t}^{FI}, \underline{t}^{FI} \mid \underline{\theta}]/\partial \underline{t}^{FI} = 0$; in turn this implies $\partial v_1[\underline{t}^{FI}, \underline{t}^{FI} \mid \overline{\theta}]/\partial \underline{t}^{FI} > 0$. Since $\gamma > 0$, a choice of $\underline{t} = \underline{t}^{FI}$ would make the left-hand side of (59) negative; concavity implies that $\partial v_1[\underline{t}, \underline{t} \mid \underline{\theta}]/\partial \underline{t}$ is decreasing in $\underline{t}$; thus, for (59) to be satisfied, we must have $\underline{t} < \underline{t}^{FI}$.

References

Bacchetta, P., and Espinosa, M.P. (1995). "Information Sharing and Tax Competition among Governments," *Journal of International Economics* 39, 103-121.

Bordignon, M., Manasse, P., and Tabellini, G. (1996). "Optimal regional Redistribution under Asymmetric Information," mimeo.

Cornes, R.C., and Silva, E.C.D. (1996a). "Transfers between Jurisdictions with Private Information: The Equity/Efficiency Trade-off," mimeo.

Cornes, R.C., and Silva, E.C.D. (1996b). "Public Good Mix in a Federation with Incomplete Information," mimeo.

Dasgupta, P.S., Hammond, P.J., and Maskin, E.S. (1979). "The Implementation of Social Choice Rules," *Review of Economic Studies* 46, 185-216.

de Crombrugghe, A., and Tulkens, H. (1990). "On Pareto Improving Tax Changes under Fiscal Competition," *Journal of Public Economics* 41, 335-350.

Gordon, R.H., and Bovenberg, A.L. (1996). "Why Is Capital So Immobile Internationally? Possible Explanations and Implications for Capital Income Taxation," *American Economics Review* 86, 1057-1075.

Gordon, R.H., and Wilson, J.D. (1986). "An Examination of Multijurisdictional Corporate Income Taxation under Formula Apportionment," *Econometrica* 54, 1357-1373.

Huizinga, L., and Nielsen, S.B. (1997). "The Taxation of Interest in Europe: A Minimum Withholding Tax?" mimeo, Economic Policy Research Unit, University of Copenhagen.

Kanbur, R., and Keen, M. (1993). "Jeux Sans Frontières: Tax Competition when Countries Differ in Size," *American Economic Review* 83, 877-892.

Keen, M. (1987). "Welfare Effects of Commodity Tax Harmonisation," *Journal of Public Economics* 33, 107-114.

Keen, M. (1989). "Pareto-Improving Indirect Tax Harmonization," *European Economic Review* 33, 1-12.

Keen, M. (1993). "The Welfare Effects of Tax Co-ordination in the European Community," *Fiscal Studies* 14, 15-36.

Laffont, J.J., and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation*, Boston: MIT Press.

Lahiri, S., and Raimondos-Møller, P. (1996). "Public Good Provision and the Welfare Effects of Indirect Tax Harmonisation," forthcoming in the *Journal of Public Economics*.

Lockwood, B. (1993). "Commodity Tax Competition under Destination and Origin Principles," *Journal of Public Economics* 52, 141-162.

Mintz, J., and Tulkens, H. (1986). "Commodity Tax Competition between Members States of a Federation: Equilibrium and Efficiency," *Journal of Public Economics* 29, 133-172.

Mirrlees, J.A. (1971). "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies* 38, 175-208.

Raff, H., and Wilson, J.D. (1997). "Income Redistribution with Well-Informed Local Governments," mimeo, Department of Economics, Indiana University.

Turunen-Red, A.H., and Woodland, A.D. (1990). "Multilateral Reform of Domestic Taxes," *Oxford Economic Papers* 42, 160-186.

Zodrow, R.G., and Mieszkowski, P. (1986). "Pigou, Tiebout, Property Taxation and the Underprovision of Local Public Goods," *Journal of Urban Economics* 19, 356-370.