Capital Taxation in a Fiscal Union – Implications of Simultaneous Horizontal and Decentralized Leadership*

Tomas Sjögren**
Department of Economics
Umeå School of Business and Economics
Umeå University, SE - 901 87 Umeå, Sweden

Abstract

This article concerns capital taxation and public good provision in a two-layer fiscal union where the federal government uses lump-sum transfers to redistribute resources between two local jurisdictions, and where each local government uses a capital tax and a lump-sum tax to finance the provision of a local public good. The novelty is to allow for simultaneous horizontal and decentralized leadership (double leadership) which means that one of the local governments is able to exercise Stackelberg leadership both vis-a-vis the other local government and vis-a-vis the federal government. Among the results it is shown that the capital tax becomes redundant as a policy instrument for the double leader if the other state government acts as a decentralized leader vis-a-vis the federal government. If, instead, the other state government does not exercise leadership vis-a-vis the federal government then the double leader will implement a capital tax which is allocatively efficient from the perspective of the fiscal union as a whole. It is also shown that double leadership exacerbates the under-taxation inefficiency that earlier research has shown exists in a fiscal union with decentralized leadership.

Keywords: Federalism, capital taxation, commitment, leadership.
JEL Classification: H10, H21, H77.

* Research grants from the Swedish Research Council (ref 421-2010-1420) are gratefully acknowledged.

** E-mail: tomas.sjogren@econ.umu.se. Phone: +46 (0)90 – 786 99 94.

1. Introduction

An important topic in public economics regards the taxation of mobile capital. Following the seminal work of Zodrow and Mieszkowski (1986), several studies have shown that if jurisdictions compete for mobile capital, then the uncoordinated Nash equilibrium with simultaneous moves may result in undertaxation of capital.[1] This inefficiency can also be related to the externalities that the jurisdictions impose on each other if they ignore that their taxes may affect the tax revenue (fiscal externality) and the interest rate (pecuniary externality) in the other jurisdictions.[2]

The theoretical literature on the taxation of mobile capital has been extended in different ways. One has been to challenge the assumption that state governments move simultaneously. Already Schelling (1960) pointed out that the viability of an equilibrium with simultaneous moves is dubious as soon as countries' commitment is considered. A similar concern was raised by Keen and Konrad (2013) who, in a survey of theoretical models of tax competition, argued that "… timing is an essential aspect in strategic games".[3] Kempf and Rota-Graziozi (henceforth KRG) (2010) were the first to endogenize the timing of the jurisdictions´ decisions. KRG follow the literature on duopoly games[4] and consider a two stage timing game where the jurisdictions in the first stage commit to move early or late and in the second stage choose the capital taxes. In the following, the first-mover in this type of game will be referred to as a *horizontal leader*. KRG show that the Subgame Perfect Equilibria (SPE) are the two Stackelberg outcomes, implying that moving simultaneously is not commitment robust for competing governments in their model. Since there are two possible SPE, there is a coordination problem. To solve this problem, KRG consider a quadratic specification of the production function and then use the concepts of Pareto-dominance and risk-dominance as selection criteria to determine which jurisdiction will be the first-mover. In this context, it is shown that the SPE where the less productive jurisdiction leads risk-dominates the other SPE. A key assumption in KRG´s model is that capital is owned by absentee owners and Ogawa (2013) shows that the simultaneous-move outcome (i.e. the Nash equilibrium) will prevail as an SPE if the capital instead is owned by the residents in the jurisdictions. Eichner (2014), in turn, extends the horizontal leadership

[1] See e.g. Bucovetsky and Wilson (1991), Keen and Marchand (1997), and Wilson (1999).
[2] Wildasin (1989) and DePater and Myers (1994).
[3] Gordon (1992) suggested that the United States may have been large and influential enough to have played the role of a Stackelberg leader. In an empirical study using data from 1968 to 2008, Altshuler and Goodspeed (2015) test the leadership role of the United States, the United Kingdom and Germany in tax competition and they do find support for a US leadership role.
[4] See Hamilton and Slutsky (1990).

game by analyzing the outcome when the state governments use the tax revenue from capital taxation to provide local public goods.

Another extension of the theoretical literature on the taxation of mobile capital has been to consider tax competition when the competing states are members of a fiscal union.[5] Köthenbürger (2004) argues that tax competition within a fiscal union is likely to differ compared with tax competion between sovereign states. One reason is that political and legal barriers are normally removed within an fiscal union. As a consequence, intra-union capital mobility tends to be higher than inter-union capital mobility, thereby magnifying concerns for tax competition in the former context. Another reason is that since member states in a fiscal union are linked by a common federal transfer system, the latter mechanism may significantly affect the incentives to engage in tax competition. The reason is that states within a fiscal union may rationally anticipate how the federal level will respond to policy changes at the state level. If the federal government treats the tax policies implemented by the state governments as given, then the policy choices within the fiscal union can be viewed as a sequential game where the state governments choose their taxes in the first stage while the federal government chooses transfers in the second stage. This sequence of events is referred to as *decentralized leadership*[6] and several studies have pointed out that this seems to be an appropriate characterization of the European Union (EU).[7] Köthenbürger (2004) analyzes tax competition under decentralized leadership and he considers a two-layer fiscal union with mobile capital where each member state chooses its capital tax to finance the provision of a local public good while the federal government uses lump-sum transfers to redistribute resources between the member states. A key result is that when the member states are identical, then ex-post federal policy neutralizes horizontal fiscal and pecuniary externalities (i.e. capital mobility itself does not lead to under-taxation of capital). Another result is that the federal redistribution mechanism will introduce a new source of inefficiency which leads to under-provision of the local public good.[8]

---

[5] In several countries with a federal structure such as Canada, Germany and USA, the taxing powers are partially delegated to the lower level governments. This makes it possible for them to compete for mobile capital.

[6] Early studies on the implications of decentralized leadership in a fiscal federal framework are Silva and Caplan (1997), Caplan and Silva (1999), Caplan et al (2000) and Aronsson et al (2006). These studies are, respectively, concerned with transboundary pollution, acid rain, non-cooperative provision of pure public goods by regional governments and environmental policy.

[7] In the EU, the Commission may be thought of as representing "the federal level." The political power of the Commission stems from its right to initiate legislation. However, in order to become EU law, its proposals must pass the Council of Ministers. In the Council of Ministers, each member acts on behalf of his/her national government. Therefore, given this decision structure, commitments at the national level are likely to affect the policies decided upon by the EU.

[8] The source of this inefficiency is that each local state government (all else equal) recognizes that one extra dollar raised in local tax revenue will induce the federal government to redistribute more resources to the other member

These earlier studies have significantly improved our understanding of the incentives, and the resulting welfare implications, underlying the taxation of mobile capital under horizontal and decentralized leadership, respectively. There are, however, no previous studies which analyze horizontal and decentralized leadership simultaneously in a unified context. A motivation for combining these approaches is that several earlier studies have pointed out that decentralized leadership seems to be an appropriate characterization of the European Union.[9] At the same time, there are also studies which have considered the possibility that some of the large EU countries (such as Germany and the United Kingdom) may be able to exercise horizontal leadership.[10] This suggests that some countries may indeed be in a situation where they are able to simultaneously exercise horizontal and decentralized leadership.

Given the arguments presented above, the main purpose of this paper is to analyze the implications for capital tax policy and welfare when one of the local governments in a fiscal union exercises horizontal leadership vis-à-vis the other local governments at the same time as that local government also exercises decentralized leadership vis-à-vis the federal government. This will be referred to as *double leadership*. To assess how the tax policy under double leadership differs both from the tax policy implemented under (i) horizontal leadership (without decentralized leadership) and from the tax policy implemented under (ii) decentralized leadership (without horizontal leadership), we will also characterize the policies implemented in cases (i) and (ii). These characterizations will provide points of reference against which we may compare the results derived under double leadership. These characterizations will also themselves provide novel results. To be more specific, in Köthenbürger´s (2004) analysis of capital taxation under decentralized leadership, it is assumed that all local jurisdictions within the fiscal union are identical. However, it is also conceivable that the local jurisdictions within a fiscal union differ in terms of their abilities to commit to a given policy. Therefore a distinction will be made between a fiscal union where

(a) *all* local governments exercise decentralized leadership vis-à-vis the federal government and a fiscal union where
(b) only *some* of the local governments exercise decentralized leadership while the other local governments treat the federal decision variables as exogenous.

states. This provides an incentive to raise less tax revenue than otherwise which leads to under-provision of the local public good.

[9] See e.g. Köthenbürger (2004) and Aronsson et al (2006).

[10] See e.g. Gordon (1992).

It turns out that the equilibria may differ considerably depending on whether we consider a game of type (a) or of type (b).

To analyze the issues discussed above, we set up a model of a two-layered fiscal union with two member states. The local government in each member state can use a capital tax and a lump-sum tax[11] to finance the provision of a local public good while the federal government uses lump-sum transfers to redistribute resources between the two jurisdictions. In line with Hoyt (2001), the federal government treats the taxes implemented by the local governments as exogenous and we first characterize the outcome in a non-cooperative Nash equilibrium where each local jurisdiction treats the choices made by the other local government, as well as the choices made by the federal government, as exogenous. When a local government behaves like this, it will be referred to as a *pure follower*. Thereafter we characterize the tax and expenditure policy when one of the local jurisdictions is able to act as a first-mover vis-à-vis the other local jurisdiction (horizontal leadership), followed by a comparison between the levels of welfare in the non-cooperative Nash equilibrium and in the Stackelberg equilibrium with horizontal leadership.

After that, we modify the model to allow for decentralized leadership. We begin by characterizing the equilibrium when each local jurisdiction acts as a Stackelberg leader (i.e. is a first-mover) vis-à-vis the federal government but where each local jurisdiction treats the policy implemented by the other local jurisdiction as exogenous. Then we modify the game by allowing one of the local jurisdictions to exercise double leadership. Finally, we characterize the outcome when only one of the local jurisdictions is able to exercise decentralized leadership while the other jurisdiction acts as a pure follower. Also here we characterize the non-cooperative Nash equilibrium, followed by a characterization of the Stackelberg equilibrium under double leadership.

The approach described above allows us to derive several novel results. One key result is that the capital tax policy implemented by the double leader will depend on whether or not the other local government is able to act as a first-mover vis-à-vis the federal government. If the other local government cannot exercise decentralized leadership, then the double leader will implement a capital tax which produces a first-best allocation of capital within the fiscal union. If the other local government instead is able to act as a first-mover vis-à-vis the federal government, then the capital tax will be redundant as a policy instrument for the double leader. Another result is that a local government perceives that the marginal cost of public funds is

[11] The ability to use a lump-sum tax means that the capital tax implemented by a local jurisdiction is a consequence of optimization instead of some arbitrary restriction on the available tax instruments.

larger when it exercises double leadership compared with when the local government only exercises pure decentralized leadership. As such, the under-taxation inefficiency that Köthenbürger (2004) showed will arise under pure decentralized leadership is exacerbated under double leadership.

In this paper, we also derive novel results regarding the non-cooperative Nash equilibrium under decentralized leadership. One key result is that each local jurisdiction´s reaction function w.r.t. the capital tax has a slope equal to one in the capital tax space. This means that the best-reply correspondence is not a contraction and implies that a unique Nash equilibrium may not exist. It is shown that a Nash equilibrium will exist either (i) if the local jurisdictions are identical and the resulting equilibrium is symmetric (this corresponds to the equilibrium characterized by Köthenbürger 2004) or (ii) if the production functions in the jurisdictions are quadratic and have identical second derivatives.

The outline of the paper is as follows. In Section 2 we present the basic model. In Section 3 we characterize the Nash game in a fiscal union where the two local governments act as pure followers followed by a characterization of the equilibrium when one local government is able to exercise horizontal leadership. In Section 4, we characterize the outcome when the local governments exercise decentralized leadership. Here we both characterize the non-cooperative Nash equilibrium under decentralized leadership and the Stackelberg equilibrium under double leadership. In Section 5, we characterize the outcome when one of the local governments acts as a pure follower while the other government either exercises pure decentralized leadership or exercises double leadership. The paper is concluded in Section 6.

2. The Basic Model

Consider a fiscal union made up of two member states and a federal government. Each state is made up of a representative household and a representative firm. The preferences of the household living in state $n = i, j$ are represented by the utility function

$$U_n = u(c_n) + \phi(g_n) \tag{1}$$

where c_n is private consumption and g_n is a public good which is provided by the fiscal authority in state n. The sub-utility functions $u(c_n)$ and $\phi(g_n)$ are increasing and strictly concave in their respective arguments. The household´s budget constraint is given by $c_n = \pi_n + r\bar{k}_n - T_n$, where π_n is (profit) income generated by a fixed factor (labor or land), r is the interest rate, $\bar{k}_n$ is a fixed endowment of a mobile factor of production (capital) and T_n is a lump-sum tax levied by the government in state n. In each state the endowment of the fixed factor is normalized to one and a single homogenous good is produced using a constant returns to scale production function $f_n(k_n, \bar{l}_n)$, where k_n is capital and $\bar{l}_n = 1$ is the fixed factor. The homogenous good can be used on a one-to-one basis for private and public consumption and since the endowment of the fixed factor is normalized to one, the production function can be written as $f_n(k_n)$, where $f_n''(k_n) < 0 < f_n'(k_n)$ and where $f_n' = df_n/dk_n$. The factor income that accrues to the fixed factor is given by $\pi_n = f_n(k_n) - (r + t_n)k_n$, where t_n is a per unit tax on capital levied by the government in state n. A profit maximizing input choice is determined by the first-order condition $f_n'(k_n) = r + t_n$ and since the capital is perfectly mobile between the two states, equilibrium on the capital market features

$$\Delta = f_i'(k_i) - f_j'(k_j) \tag{2}$$

$$\bar{k}_i + \bar{k}_j = k_i + k_j \tag{3}$$

where $\Delta = t_i - t_j$. Since an economically efficient allocation of capital implies $f_i'(k_i) = f_j'(k_j)$, it follows that the absolute value of Δ can be used as a measure of the tax distortion within the fiscal union. Equations (2) and (3) implicitly define the capital allocation as a function of Δ, i.e. $k_i(\Delta) = k_i(t_i - t_j)$ and $k_j(\Delta) = k_j(t_i - t_j)$, while the firm´s first-order condition for capital can be used to define the interest rate as a function of the capital taxes; $r(t_i, t_j) = f_i'\left(k_i(t_i - t_j)\right) - t_i$. The comparative static properties of these functions are

$$\frac{\partial k_i}{\partial \Delta} = \frac{\partial k_i}{\partial t_i} = \frac{\partial k_j}{\partial t_j} = -\frac{\partial k_j}{\partial t_i} = -\frac{\partial k_i}{\partial t_j} = \frac{1}{f_i''(k_i) + f_j''(k_j)}, \qquad \frac{\partial r}{\partial t_i} = -\frac{f_j''(k_j)}{f_i''(k_i) + f_j''(k_j)} \tag{4}$$

The public sector is modelled as a two-layered federal system where the federal government uses lump-sum transfers to redistribute resources between the member states. If we let R_n denote the net transfer of state n to the federal government, the budget constraint facing state government n can be written as $g_n = t_n k_n + T_n - R_n$. The objective of government n is to choose t_n and T_n to maximize the welfare function $W_n = U_n$ while the objective of the federal government is to redistribute resources to maximize the aggregate welfare; $W = U_i + U_j$. The federal budget constraint is given by $R_i + R_j = 0$ and the federal government treats the taxes chosen by the state governments as exogenous. By substituting the state governments´ budget constraints into W and maximizing the resulting expression subject to the federal budget constraint, it is straightforward to show that the optimal federal policy will feature complete redistribution in the sense that $\phi'(g_i) = \phi'(g_j)$, where $\phi' = d\phi/dg$. This equation and the federal budget constraint together define the net transfers of states *i* and *j* to the federal government as functions of the local governments´ decision variables; $R_i = R_i(t_i, T_i, t_j, T_j)$ and $R_j = R_j(t_i, T_i, t_j, T_j)$. These functions will be referred to as federal reaction functions and the comparative static properties of these functions are

$$\frac{\partial R_i}{\partial T_i} = -\frac{\partial R_j}{\partial T_i} = \frac{\phi_i''}{\phi_i''+\phi_j''} = \frac{1}{2}, \qquad \frac{\partial R_i}{\partial t_i} = -\frac{\partial R_j}{\partial t_i} = \frac{\phi_i''\left(k_i + t_i\frac{\partial k_i}{\partial t_i}\right) - \phi_j'' t_j \frac{\partial k_j}{\partial t_i}}{\phi_i''+\phi_j''} \tag{5}$$

where $\phi_n'' = \phi''(g_n)$ and $\phi''(g_i) = \phi''(g_j)$.[12]

Let us now turn to the tax and expenditure problem facing the state governments and we begin by characterizing the outcome when each state government treats the choices made by the federal government as exogenous.

3. A Fiscal Union Without Decentralized Leadership

In this section we consider a fiscal union where the state governments treat the federal transfers (R_i, R_j) as exogenous. In Subsection 3.1 we characterize the non-cooperative Nash equilibrium when each state government acts as a pure follower and chooses its decision variables while treating the decision variables of the other state government as exogenous. Since we here replicate some well-known results from earlier studies, the analysis in this subsection will serve as a point of reference against which we can compare the results to be derived in subsequent parts of the paper. In Subsection 3.2 we extend the analysis by characterizing the outcome when government *i* is able to exercise horizontal leadership vis-à-vis government *j*.

3.1. The Non-Cooperative Nash Equilibrium

Each state government accounts for the effects on capital demand and on the interest rate via equations (2) - (4). By substituting the private and the public budget constraints into equation (1), and by using the definition of the profit, the objective function of government i (the problem facing government *j* is analogous) can be written as

$$W_i(t_i, T_i, t_j) = u\left(f_i(k_i) - f_i'(k_i)k_i + r\bar{k}_i - T_i\right) + \phi(t_i k_i + T_i - R_i) \tag{6}$$

[12] Since the agents are assumed to have identical preferences, and since federal redistribution features $\phi'(g_i) = \phi'(g_j)$, it follows that $\phi''(g_i) = \phi''(g_j)$.

where $k_i = k_i(t_i - t_j)$ and $r = r(t_i, t_j)$. Let us begin by defining $MCPF_i = \phi'(g_i)/u'(c_i)$ to be the marginal cost of public funds in state *i*. Then, maximizing $W_i(t_i, T_i, t_j)$ w.r.t. T_i produces $MCPF_i = 1$, which implies that the provision of the local public good will be efficient. By using this result in the first-order condition for t_i, it can be shown that the capital tax implemented by the government in state *i* is implicitly determined by[13]

$$t_i = -(\bar{k}_i - k_i)\frac{\partial r/\partial t_i}{\partial k_i/\partial t_i} = (\bar{k}_i - k_i)f_j'' \quad (7)$$

where we in the second step have used the comparative static properties in (4). The corresponding tax formula for government *j* is $t_j = (\bar{k}_j - k_j)f_i''$. Equation (7) shows that the only reason for implementing a non-zero capital tax is the pecuniary motive to manipulate the terms of trade (i.e. to influence the interest rate).[14] Analogous results can be found in Peralta and van Ypersele (2005), Itaya et al (2008) and Ogawa (2013).

Equation (7) implicitly defines t_i as a (reaction) function of t_j, $t_i^{PF}(t_j)$, where the super-index "*PF*" highlights that this is the reaction function when the government acts as a Pure Follower. By differentiating equation (7) w.r.t t_i and t_j, and by using (3) and (4), we obtain the following expression for the slope of the reaction function

$$\frac{dt_i^{PF}}{dt_j} = \frac{f_j'' - (\bar{k}_j - k_j)f_j'''}{f_i'' + 2f_j'' - (\bar{k}_j - k_j)f_j'''} \quad (8)$$

Equation (8) implies that the slope of the pure follower´s reaction function will be situated in the interval $0 < dt_i^{PF}/dt_j < 1$ as long as the condition $f_j'' - (\bar{k}_j - k_j)f_j''' < 0$ is satisfied. A positively sloped reaction function is known as strategic complementarity[15] while the property $dt_i^{PF}/dt_j < 1$ implies that the best-reply correspondence is a contraction. The latter property is needed to guarantee the existence and uniqueness of a Nash equilibrium between two pure followers[16] where $t_i^{NE} = t_i(t_j^{NE})$ and $t_j^{NE} = t_j(t_i^{NE})$ hold simultaneously (the super-index "*NE*" denotes an equilibrium value in the Nash Equilibrium). Since the existence and uniqueness of the Nash equilibrium rests on the assumption that $f_j'' - (\bar{k}_j - k_j)f_j''' < 0$ holds,

[13] See the Appendix.
[14] Since lump-sum taxes are available at the state level, the fiscal motive is not present in this model.
[15] Strategic complementarity is a standard assumption in tax competition models. See, for example, Bucovetsky (2009) and Konrad and Schjeldrup (1999).
[16] See Vives (1999) and Kempf and Rota-Graziosi (2010).

this inequality will be referred to as the *existence condition*. Finally, we note that the other government´s reaction function, $t_j^{PF}(t_i)$, can be defined analogously.

If the production functions and the initial allocations of capital in the two states are identical, then $f_i'(\bar{k}_i) = f_j'(\bar{k}_j)$. In this situation $(\bar{k}_i - k_i) = (\bar{k}_j - k_j) = 0$, in which case equation (7) implies that the capital taxes will be zero in the resulting symmetric equilibrium. In this case there is no allocative distortion on the capital market and the outcome will be efficient. If, on the other hand, $f_i'(\bar{k}_i) \neq f_j'(\bar{k}_j)$ holds for the initial allocations (this may either be because the initial allocations differ and/or because the production functions are not identical), then one state will become a net importer, and the other state will become a net exporter, of capital in the resulting Nash equilibrium, i.e. $(\bar{k}_i - k_i^{NE}) = -(\bar{k}_j - k_j^{NE}) \neq 0$. By using this property in (7), it follows that if state *i* is the net importer capital $(\bar{k}_i < k_i^{NE})$ then government *i* implements a positive tax on capital to improve its terms of trade (i.e. to reduce r) while government *j* implements a negative capital tax to push up the interest rate. As a consequence, t_i^{NE} and t_j^{NE} will have opposite signs.[17] If we use these results in equation (2), it follows that $f_i'(k_i^{NE}) - f_j'(k_j^{NE}) = t_i^{NE} - t_j^{NE} \neq 0$ which, in turn, implies that the non-symmetric Nash equilibrium will be inefficient. This finding is analogous to results derived by Ogawa (2013).

The Nash equilibrium is illustrated in Figure 1 for the case where $f_i'(\bar{k}_i) > f_j'(\bar{k}_j)$ which means that state *i* will become the net importer, and state *j* will become the net exporter, of capital in the resulting Nash Equilibrium (NE). In the Appendix it is shown that the two reaction functions intersect below the Zero Capital Export (ZCE) locus in the second quadrant in (t_j, t_i) – space, as illustrated in Figure 1 below. The ZCE locus is implicitly defined by the equation $0 = \bar{k}_i - k_i(t_i - t_j)$ and its slope is one.

[17] If the capital would be owned by absentee owners (as in Kempf and Rota-Graziosi 2010), or if the local fiscal authorities cannot not use lump-sum taxes (as in Eichner 2014), then the Nash equilibrium would feature positive capital taxes in both member states.

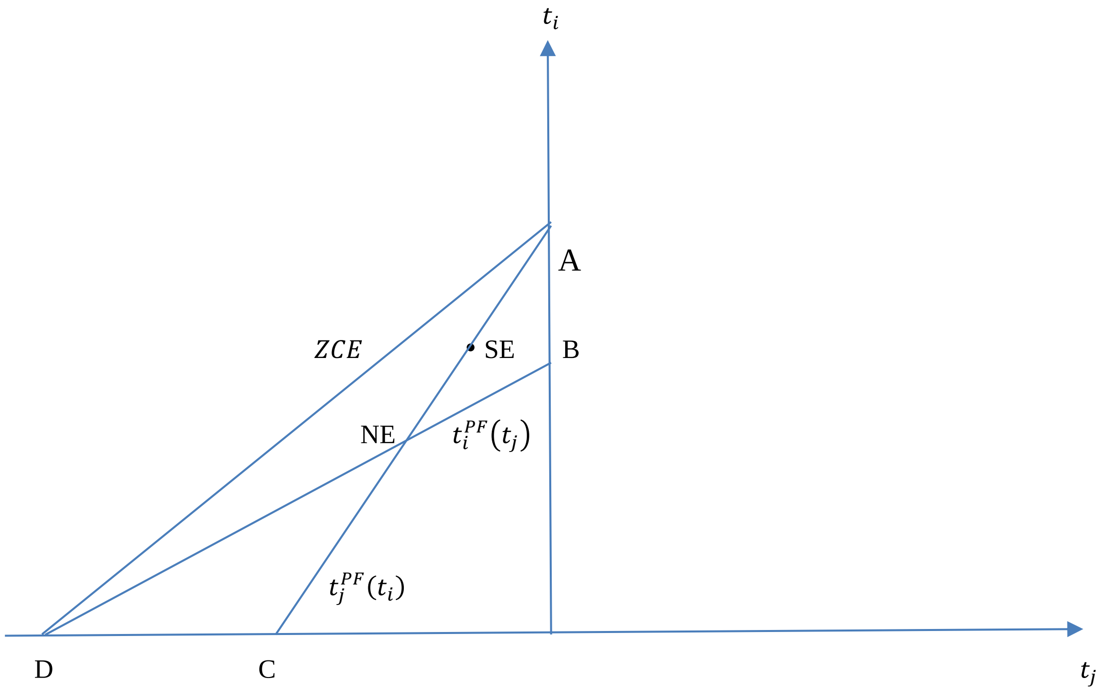

Figure 1. Nash and Stackelberg equilibria when state *i* is the net importer, and state *j* is the net exporter, of capital. The Nash equilibrium with two pure followers is point NE while point SE is the Stackelberg equilibrium when state *i* exercises horizontal leadership. A and C are the points where the $t_j^{PF}(t_i)$ function intersects with the vertical and horizontal axis, respectively, while B and D are the points where the $t_i^{PF}(t_j)$ function intersects with the vertical and horizontal axis.

3.1. Horizontal Leadership

Let us now consider the outcome when government i is able to act as a first-mover vis-à-vis government j. As a follower, government j chooses its tax policy as outlined in Section 3.1 while government i takes government j´s reaction function $t_j^{PF}(t_i)$ into account when it maximizes[18] $W_i\left(t_i, T_i, t_j^{PF}(t_i)\right)$ w.r.t. t_i and T_i. As before, the optimal choice of T_i will feature $MCPF_i = 1$ but in the Appendix it is shown that the optimal capital tax will now (implicitly) be determined by

$$t_i = (\bar{k}_i - k_i)f_j'' + (\bar{k}_i - k_i)\left[f_j'' - (\bar{k}_j - k_j)f_j'''\right] \tag{9}$$

Comparing equation (9) with equation (7) shows that as long as the existence condition is satisfied, then the horizontal leader will set the *absolute value* of its capital tax higher than in the Nash equilibrium. To explain this result, consider the case when state i is the net importer

[18] There is also a reaction function associated with the lump-sum tax in the other state but since T_j does not appear anywhere in the horizontal leader´s objective function, effects via that reaction function are redundant.

of capital ($\bar{k}_i < k_i$). As a net importer, state i would benefit from a reduction of the interest rate. Since an increase of the capital tax in the other state has a negative effect on the interest rate, and since the other state´s reaction function $t_j = t_j^{PF}(t_i)$ has a positive slope, this knowledge induces government *i* to set t_i higher in the Stackelberg equilibrium (SE) than in the NE. Furthermore, since $0 < dt_j^{PF}/dt_i < 1$, it follows that $t_i^L - t_j^F > t_i^{NE} - t_j^{NE} > 0$, where the super-indexes "*L*" and "*F*" denote the "leader" and the "follower", respectively, in the SE. This SE is illustrated in Figure 1 and the argument for why $t_i^L - t_j^F < t_i^{NE} - t_j^{NE} < 0$ when the horizontal leader instead is the net exporter of capital is analogous.

These arguments imply the following result;

Lemma 1: *As long as the existence condition $f_j'' - (\bar{k}_j - k_j)f_j''' < 0$ is satisfied, the size of the tax distortion, $|\Delta| = |t_i - t_j|$, will be larger in the Stackelberg equilibrium with horizontal leadership than in the non-cooperative Nash equilibrium with pure followers.*

Since the size of the tax distortion will be larger in the SE than in the NE, it follows that the aggregate output, $f_i(k_i) + f_j(k_j)$, will be larger in the latter equilibrium. Furthermore, since the state governments in both equilibria provide the local public goods efficiently according to the rule $u'(c_n) = \phi'(g_n)$, and since federal redistribution is efficient in the sense that $\phi'(g_i) = \phi'(g_j)$, it follows that the aggregate welfare, $W = U_i + U_j$, will be a monotonously increasing function of aggregate output. Together, these arguments imply the following result;

Proposition 1: *Consider a federal system made up of two state governments* (*i and j*) *and a federal government, where the latter uses federal transfers to achieve complete redistribution in the sense that $\phi'(g_i) = \phi'(g_j)$. If one of the state governments is able to exercise horizontal leadership, then the aggregate welfare* ($U_i + U_j$) *will be lower in the Stackelberg equilibrium than in the Nash equilibrium.*

This result is a direct consequence of the federal redistribution that takes place within the fiscal union. If there would be no federal government which redistributes resources according to the rule $\phi'(g_i) = \phi'(g_j)$, then the aggregate welfare need not be a monotonously increasing function of the aggregate output.[19] In that situation the arguments underpinning Proposition 1

[19] Consider, for example, a situation without federal redistribution where the national income and the welfare are smaller in state *i* than in state *j* in the Nash equilibrium; i.e. $f_i(k_i^{NE}) + r(\bar{k}_i - k_i^{NE}) < f_j(k_j^{NE}) + r(\bar{k}_j - k_j^{NE})$ and $W_i^{NE} < W_j^{NE}$. In this situation, the aggregate welfare may be improved if state *i* is able to exercise horizontal leadership vis-a-vis state *j* because the redistribution of resources that occurs when the economy moves from the

are no longer valid and the welfare comparison between the Nash equilibrium and the Stackelberg equilibrium would be inconclusive unless additional assumptions are made regarding e.g. the initial distribution of capital.

4. A Fiscal Union With Two Decentralized Leaders

In this section we allow the local governments to act as first-movers vis-à-vis the federal government. In Subsection 4.1, we characterize the non-cooperative Nash equilibrium between the two state governments when they act as Nash followers vis-à-vis each other. In Subsection 4.2 we extend the analysis and study the outcome when state government *i* is also able to exercise horizontal leadership vis-à-vis state government *j*.

4.1. The Non-Cooperative Nash Equilibrium

When state government *i* (the problem facing government *j* is analogous) exercises decentralized leadership, it takes the federal reaction function $R_i = R_i(t_i, T_i, t_j, T_j)$ into account when maximizing the objective function defined in equation (6). As before, the tax policy implemented by the other state government is treated as exogenously given and government *i* accounts for the effects on the capital demand and the interest rate via the equations in (2) - (4). Substituting $R_i(t_i, T_i, t_j, T_j)$ into equation (6) and maximizing w.r.t. T_i, while we also use that $\partial R_i/\partial T_i = \phi_i''/(\phi_i'' + \phi_j'') = 0.5$, produces

$$MCPF_i = \frac{\phi_i'' + \phi_j''}{\phi_j''} = 2 \tag{12}$$

The result in equation (12) is analogous to the result derived by Köthenbürger (2004) that the provision of the public good will be inefficiently low under decentralized leadership. The reason is that the state government recognizes that if it raises one extra dollar in tax revenue then the federal government will respond by increasing the net contribution that state i pays to the federal level by 50 cents. Therefore the marginal cost of raising one dollar in tax revenue exceeds one even though the state government can use a lump-sum tax to raise its tax revenue.

Turning to the optimal capital tax, it will be given by (see the Appendix)

Nash equilibrium to the Stackelberg equilibrium may improve the welfare in state *i* by so much that this welfare improvement exceeds the welfare loss in state *j*.

$$t_i = (\bar{k}_i - k_i)f_j'' - t_j \frac{\partial k_j/\partial t_i}{\partial k_i/\partial t_i} \tag{13}$$

The first term on the RHS of equation (13) is identical to, and can be given the same interpretation as, the corresponding term on the RHS of equation (7). The second term on the RHS of equation (13) is novel and is directly related to government *i*´s ability to exercise decentralized leadership. To explain why this term appears, observe that even if government *i* treats the tax instruments chosen by government j (t_j and T_j) as exogenous, government *i* nevertheless recognizes that the revenue from capital taxation in the other state, $t_j k_j$, will be influenced by t_i via the capital demand function, $k_j(t_i - t_j)$. From its perspective, government *i* therefore perceives that if $t_j \partial k_j/\partial t_i$ is positive (negative), then an increase in t_i leads to more (less) tax revenue in the other state which, all else equal, provides the federal government with an incentive to redistribute more (less) resources from state *j* to state *i*. Since this effect only appears when government *i* is able to exercise decentralized leadership, the second term on the RHS in equation (13) will be referred to as the *decentralized leadership motive* for capital taxation. Since we know from (4) that $\partial k_j/\partial t_i > 0$, it follows that if $t_j > 0$, then the decentralized leadership motive provides government *i* with an incentive to set t_i higher than otherwise but if $t_j < 0$, then the opposite argument applies. This property can be made more explicit if we use the comparative static properties in (4) to rewrite equation (13) to read

$$t_i = (\bar{k}_i - k_i)f_j'' + t_j \tag{13´}$$

This equation implicitly defines t_i as a (reaction) function of t_j, $t_i^{DL}(t_j)$, where the super-index "*DL*" refers to Decentralized Leader. The solution to the other state *j*´s maximization problem produces an analog reaction function; $t_j^{DL}(t_i)$. To derive an expression for the slope of a decentralized leader´s reaction function, we differentiate equation (13´) w.r.t. t_i and t_j. By using the comparative static properties in (4), the following result is readily available;

Proposition 2: *The slope of a decentralized leader´s reaction function is equal to one.*

This result implies that when both state governments exercise decentralized leadership, then the two states´ reaction functions cannot cross each other in (t_j, t_i) space. As a consequence, it is not possible to retrieve a unique Nash equilibrium in (t_j, t_i) space. Instead, Nash equilibria can only exist if the two reaction functions over-lap, i.e. if the two reaction functions have the same intercept and slope in (t_j, t_i) space. In this situation, any (t_j, t_i) combination along the joint line is a possible Nash equilibrium.

However, it turns out that all points along a given reaction function are associated with a unique capital allocation. To verify this claim for government i´s reaction function, we substitute $t_i^{DL}(t_j)$ into the capital demand functions; $k_i\left(t_i^{DL}(t_j) - t_j\right)$ and $k_j\left(t_i^{DL}(t_j) - t_j\right)$. If we differentiate these functions w.r.t. t_j, we obtain

$$\frac{\partial k_i}{\partial t_j} = \frac{\partial k_i}{\partial t_i^{DL}} \frac{\partial t_i^{DL}}{\partial t_j} + \frac{\partial k_i}{\partial t_j} = 0, \qquad \frac{\partial k_j}{\partial t_j} = \frac{\partial k_j}{\partial t_j} + \frac{\partial k_j}{\partial t_i^{DL}} \frac{\partial t_i^{DL}}{\partial t_j} = 0 \tag{14}$$

where we have used the comparative static properties in (4) together with $\partial t_i^{DL}/\partial t_j = 1$. The fact that the total derivatives in (14) are zero implies that the capital allocation is unchanged as we move along government i´s reaction function. Therefore, there must be a unique capital allocation, which we denote $(\hat{k}_j, \hat{k}_i)$, associated with the reaction function $t_i^{DL}(t_j)$. It also follows that the tax distortion, $\Delta = t_i(t_j) - t_j$, will be fixed at a given level (denoted $\hat{\Delta}$) as we move along $t_i^{DL}(t_j)$. To determine the levels of $\hat{k}_i$, $\hat{k}_j$ and $\hat{\Delta}$, let us first substitute equations (3) and (13´) into equation (2). This produces

$$\Omega_i(k_i) = f_i'(k_i) - f_j'\left(\bar{k}_i + \bar{k}_j - k_i\right) - \left(\bar{k}_i - k_i\right) f_j''\left(\bar{k}_i + \bar{k}_j - k_i\right) = 0 \tag{15}$$

If $f_j'' + \left(\bar{k}_i - k_i\right) f_j''' < 0$,[20] then the function $\Omega_i(k_i)$ is monotonously decreasing in k_i, in which case there exists a unique solution to $\Omega_i(k_i) = 0$ which defines $\hat{k}_i$. Conditional on $\hat{k}_i$, we can then determine $\hat{k}_j$ and $\hat{\Delta}$ from equations (3) and (2), respectively. A similar approach can be used to define an analog to equation (15) for state j, $\Omega_j(k_j) = 0$, where we let $(\check{k}_j, \check{k}_i)$ denote the capital allocation along state j´s reaction function.

The fact that each reaction function is associated with a unique capital allocation implies that if the two reaction functions over-lap in (t_j, t_i) space, then the over-lapping reaction functions will be associated with a unique Nash equilibrium in terms of the capital allocations if $\hat{k}_i = \check{k}_i$ and $\hat{k}_j = \check{k}_j$. This happens in the following cases;

[20] Since $f_j'' + \left(\bar{k}_i - k_i\right) f_j''' = f_j'' - \left(\bar{k}_j - k_j\right) f_j''' < 0$, this inequality corresponds to the existence condition.

Proposition 3: *Consider a federal system made up of two state governments and a federal government where the two state governments act as decentralized leaders.*

(a) *If the two states are identical, then there exists a unique, symmetric Nash equilibrium in terms of the capital allocations which is allocatively efficient.*

(b) *If the production functions in the two states countries are quadratic and can be written as* $f_i(k_i) = a_i k_i - 0.5bk_i^2$ *and* $f_j(k_j) = a_j k_j - 0.5bk_j^2$*, respectively, where* a_i*,* a_j *and* b *are positive constants, then there exists a unique Nash equilibrium in terms of the capital allocations but this equilibrium may not be allocatively efficient.*

Part (a) in Proposition 3 is analogous to the result derived by Köthenbürger (2004) where he shows that decentralized leadership neutralizes tax competition in an economic federation made up of identical states. Although there is no direct tax competition motive behind the capital taxes implemented by the state governments in this model (because each state government can use its lump-sum tax to raise its tax revenue) we nevertheless obtain a conclusion analogous to that of Köthenbürger because in a symmetric equilibrium where $\bar{k}_i - k_i = \bar{k}_j - k_j = 0$, the terms of trade motive underlying the capital tax vanishes. In this situation, equation (13´) reduces to $t_i = t_j$ which means that the two reaction functions over-lap. Then, the functions $\Omega_i(k_i) = 0$ and $\Omega_j(k_j) = 0$ will produce the same capital allocations because the production functions are identical and $\bar{k}_i = \bar{k}_j$. The resulting Nash equilibrium will be allocatively efficient because $\Delta = f_i'(k_i) - f_j'(k_j) = 0$.

Part (b) in Proposition 3 is novel and in the Appendix it is shown that $\Omega_i(k_i) = 0$ and $\Omega_j(k_j) = 0$ produce the same capital allocations under the conditions specified in part (b). To highlight the key mechanism behind this result, observe that it follows from equation (13´) that the reaction functions of government *i* and government *j* are implicitly determined by $t_i - t_j = (\bar{k}_i - k_i)f_j''$ and $t_i - t_j = -(\bar{k}_j - k_j)f_i''$, respectively. Since $\bar{k}_i - k_i = -(\bar{k}_j - k_j)$, the two reaction functions will be identical, and will therefore over-lap in (t_j, t_i) - space, if $f_i'' = f_j''$. This latter condition is satisfied with the quadratic production functions specified in part (b). Note that the result in part (b) holds irrespective of whether the net capital export is zero or not. If $\bar{k}_i - k_i = -(\bar{k}_j - k_j) \neq 0$, then it follows from equation (13´) that $\Delta = t_i - t_j \neq 0$ which means that the resulting capital allocation will be inefficient.

4.2 Horizontal Leadership

Let us now consider the maximization problem facing state government i when it also exercises horizontal leadership vis-à-vis the other state government while we retain the assumption that both state governments exercise decentralized leadership vis-à-vis the federal government. This means that government *i* now exercises double leadership and that the decision-making process within the fiscal union is as follows:

1. In the first step, state government *i* chooses its tax policy while it simultaneously recognizes the federal reaction functions and the reaction functions of the other state government.
2. In the second step, government *j* chooses its tax policy while it recognizes the federal reaction functions but where it treats the taxes implemented by government *i* as exogenous.
3. In the final step, the federal government determines the federal transfer payments conditional on the taxes implemented by the two state governments.

This sequence of events implies that the federal government behaves as described in Section 2 while government *j* behaves as described in Section 4.1. As for government *i*, it maximizes the welfare function defined in equation (6) subject to the federal reaction functions $R_i = R_i(t_i, T_i, t_j, T_j)$ and $R_j = R_j(t_i, T_i, t_j, T_j)$, and subject to government *j*´s reaction function $t_j^{DL}(t_i)$. Government *i* also takes into account that the lump-sum tax levied in the other state, T_j, will depend on t_i and T_i via the functions $r(t_i, t_j)$, $k_j(t_i - t_j)$ and $R_j(t_i, T_i, t_j, T_j)$. Government *i*´s maximization problem is solved in the Appendix and let us begin by looking at the optimal choice of T_i which indirectly determines the marginal cost of public funds in state *i*. In the Appendix, we derive the following result;

Proposition 4: *If state government i exercises double leadership at the same time as the other state government exercises decentralized leadership, then the marginal cost of public funds facing the double leader will be given by*

$$MCPF_i = \frac{\phi_i'' + \phi_j''}{\phi_j''} + \frac{1}{MCPF_j}\frac{\phi_i''}{u_j''} > 2 \tag{16}$$

Proposition 4 shows that the marginal cost of public funds will be larger when government *i* exercises double leadership compared with when government *i* only exercises decentralized leadership (in which case $MCPF_i = 2$). This means that the undertaxation inefficiency (and hence the underprovision of the local public good) is exacerbated under double leadership. This can be explained as follows. Recall that when government *i* exercises decentralized leadership, then the government recognizes that 50 cents of one extra dollar raised in tax revenue will be collected by the federal government and transferred to the other state. Since government *i* under double leadership also takes into account the response of the other state government, government *i* will now recognize that the transfer of 50 cents to state *j* makes it possible for government *j* to reduce its lump-sum tax T_j. This reduction of tax revenue in state *j* will induce the federal government to redistribute additional resources from state *i* to state *j*, i.e. R_i will be increased further. This feed-back mechanism (which is not recognized under pure decentralized leadership) introduces an additional cost of raising tax revenue in state *i*, which is captured by the second term on the RHS of equation (16), and explains why $MCPF_i > 2$ under double leadership.

Let us now turn to the capital tax implemented under double leadership. In the Appendix we derive the following result;

Proposition 5: *If state government i exercises double leadership at the same time as the other state government exercises decentralized leadership, then any level of* t_i *satisfies government i´s first-order condition for the capital tax.*

Proposition 5 implies that the capital tax is effectively redundant as a policy instrument for the double leader when it exercises leadership vis-à-vis a state government which itself is able to exercise decentralized leadership. To explain this result, let us first recall that under pure decentralized leadership, government *i* has two motives for implementing a nonzero capital tax (see equation 13); the pecuniary motive and the decentralized leadership motive. Both these motives vanish when government *i* exercises double leadership.

Let us first explain why the decentralized leadership motive vanishes. This is a consequence of the result derived in Subsection 4.1 where it was shown that there is a unique capital allocation $(\breve{k}_j, \breve{k}_i)$ associated with the other government´s reaction function, $t_j^{DL}(t_i)$. This means that regardless of what level of t_i government *i* implements, the capital tax base in the other state, $\breve{k}_j$, will be constant.[21] This is recognized by government *i* when it exercises double

[21] Recall that the decentralized leadership motive arises because government *i* perceives that it can influence the capital tax base in the other state.

leadership[22] and implies that the decentralized leadership motive behind the capital tax vanishes.

To explain why the pecuniary motive behind capital taxation vanishes, observe first that since government *i* recognizes that the capital allocation will be fixed at $(\check{k}_j, \check{k}_i)$, government *i* will de facto treat the capital tax as a lump-sum tax. As such, t_i and T_i will be equivalent from a pure tax revenue perspective. Note, however, that t_i and T_i are not equivalent in terms of the ability to influence the interest rate because a change in T_i leaves the interest rate unchanged while a change in t_i affects the interest rate, $r\left(t_i, t_j^{DL}(t_i)\right)$, according to

$$\frac{\partial r}{\partial t_i} = \frac{\partial r}{\partial t_i} + \frac{\partial r}{\partial t_j^{DL}} \frac{\partial t_j^{DL}}{\partial t_i} = -1, \tag{17}$$

Equation (17) shows that t_i is an effective instrument to influence the interest rate also in this scenario. This result suggests (wrongly, as it turns out) that if state *i*, for example, would be a net importer of capital $(\bar{k}_i - k_i < 0)$, then the pecuniary motive would provide government *i* with an incentive to implement a higher capital tax than otherwise in order to reduce the interest rate, i.e. to improve the terms of trade of state *i*. However, an improvement in the terms of trade for state *i* is mirrored by a reduction in the terms of trade for the other state.[23] This provides the federal government with an incentive to compensate state *j* by redistributing resources from state *i* to state *j*. This "terms-of-trade" induced federal transfer response exactly offsets the benefit for state *i* of influencing the interest rate. Since government *i* recognizes this federal policy response when it exercises double leadership, the pecuniary motive vanishes.

Let us finally make a comparison between the level of welfare in the non-cooperative Nash equilibrium described in Subsection 4.1 and the level of welfare in the Stackelberg equilibrium with double leadership. Proposition 6 summarizes;

Proposition 6: *The aggregate welfare will be lower in the Stackelberg equilibrium with double leadership than in the Nash equilibrium with two decentralized leaders.*

To prove this claim, observe first that the aggregate output within the fiscal union is the same in the Nash equilibrium as in the Stackelberg equilibrium. The reason is that the equilibrium taxes in both cases are situated somewhere along state *j*´s reaction function, and from the

[22] This is not recognized under pure decentralized leadership because then the other state´s reaction functions are not incorporated into government *i*´s maximization problem.

[23] More specifically, the net national income in the other state is given by $f_j(k_j) + (\bar{k}_j - k_j)r - R_j$. Therefore, if $\bar{k}_j - k_j > 0$, it follows that a reduction of r has a negative effect on the net national income in state *j*.

analysis above we know that the capital allocation, and hence the aggregate output, will be the same at all points along state *j*´s reaction function. Observe also that an efficient distribution of the aggregate output is attained if the resources are allocated so as to satisfy $\phi'(g_i) = \phi'(g_j)$, $u'(c_i) = \phi'(g_i)$ and $u'(c_j) = \phi'(g_j)$ where the latter two conditions imply that $MCPF_i = MCPF_j = 1$. In the scenarios considered here, the federal transfer mechanism ensures that the efficiency condition $\phi'(g_i) = \phi'(g_j)$ always holds, but we know from the analysis above that the marginal costs of public funds will exceed one in both equilibria. From equation (12) we know that $MCPF_i = MCPF_j = 2$ in the Nash equilibrium with two decentralized leaders, and from the analysis in this part we know that $MCPF_i > MCPF_j = 2$ in the Stackelberg equilirium with double leadership. Hence, the distributional inefficiency will be larger, and the welfare will be lower, in the Stackelberg equilibrium with double leadership than in the corresponding Nash equilibrium.

5. A Fiscal Union With One Decentralized Leader

In this part we modify the model in Section 4 by assuming that government i is able to exercise decentralized leadership vis-à-vis the federal government while the other government *j* acts as a pure follower. Also here we will characterize two equilibria; a non-cooperative Nash equilibrium between the two state governments where they treat each other´s taxes as exogenous and a Stackelberg equilibrium where government *i* is able to exercise double leadership.

5.1. The Non-Cooperative Nash Equilibrium

From Subsection 3.1 we know that the pure follower (government *j*) implements a policy where the capital tax and the lump-sum tax are set so as to satisfy $t_j = (\bar{k}_j - k_j)f_i''$ and $MCPF_j = 1$, respectively, while we from Subsection 4.1 know that the decentralized leader (government *i*) implements a policy which satisfies $t_i = (\bar{k}_i - k_i)f_j'' + t_j$ and $MCPF_j = 2$. We also know that the slope of government *j*´s reaction function is situated in the interval $0 < dt_j^{PF} dt_i < 1$ while the slope of the decentralized leader´s reaction function is equal to one. Figure 2 depicts this Nash equilibrium for the case where $f_i'(\bar{k}_i) > f_j'(\bar{k}_j)$, which means that state *i* becomes the net importer of capital.

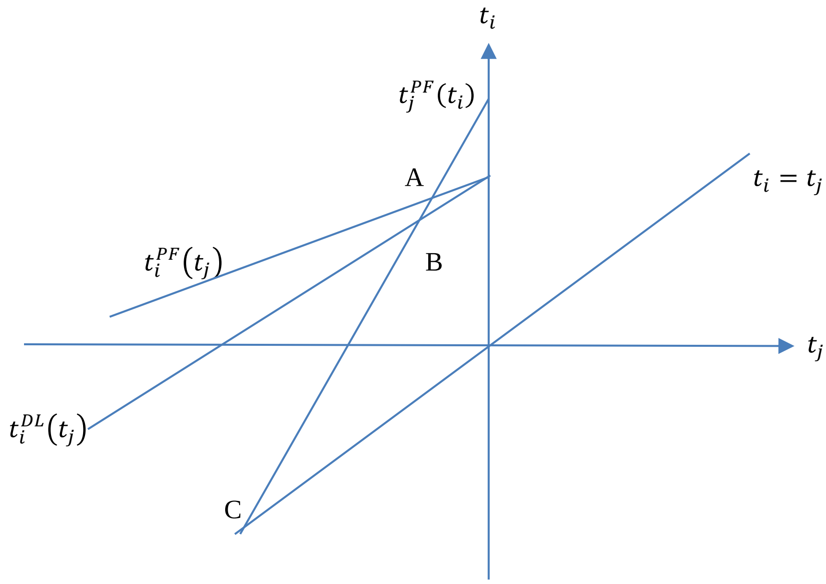

Figure 2. Nash and Stackelberg equilibria when state *i* is the net importer of capital and state *j* is the net exporter of capital. The Nash equilibrium between two pure followers is point A while point B is the Nash equilibrium when state *i* exercises a decentralized leadership while state *j* is a pure follower. Point C is the Stackelberg equilibrium when state *i* exercises double leadership vis-à-vis the pure follower.

Figure 2 depicts government *i*´s reaction function both when government *i* acts as a pure follower, $t_i^{PF}(t_j)$, and when government *i* exercises decentralized leadership, $t_i^{DL}(t_j)$. These two functions intersect on the vertical axis[24] and since $t_i^{DL}(t_j)$ has a steeper slope than $t_i^{PF}(t_j)$, it follows that $t_i^{DL}(t_j)$ will be situated below $t_i^{PF}(t_j)$ in the second quadrant in (t_j, t_i) – space. Figure 2 also depicts government *j*´s reaction function, $t_j^{PF}(t_i)$, and the *Decentralized Leader – Pure Follower* Nash equilibrium[25] (point B) will feature lower capital taxes, and a smaller tax distortion,[26] than the *Pure Follower – Pure Follower* Nash equilibrium (point A).

[24] When government *i* acts as a pure follower, then the capital tax is determined by $t_i = (\bar{k}_i - k_i)f_j''$ but when government *i* acts as a decentralized leader, then the capital tax is instead determined by $t_i = (\bar{k}_i - k_i)f_j'' + t_j$. These two functions coincide when $t_j = 0$.

[25] Observe that a necessary condition for the existence of this Nash equilibrium is that $t_i^{DL}(t_j)$ crosses the horizontal axis to the left of $t_j^{PF}(t_i)$.

[26] The move from point A to B implies a reduction of t_i by an amount that we denote by Δt_i. The corresponding change in t_j is then given by $\Delta t_j = \frac{dt_j^{PF}}{dt_i}\Delta t_i$. Since $dt_j^{PF}/dt_i < 1$, it follows that $|\Delta t_j| < |\Delta t_i|$. Then, since $t_i > 0$ and $t_j < 0$ holds at point A and at point B, it follows that the tax distortion, $\Delta = t_i - t_j$, will be smaller at point B than at point A.

5.2 Double Leadership

The maximization problem under double leadership is analogous to the corresponding maximization problem defined in Subsection 4.2 except that (i) the other state government´s reaction function is now implicitly determined by equation (7) (instead of equation 13´) and (ii) the marginal cost of public funds in the other state is now given by $MCPF_j = 1$ (instead of $MCPF_j = 2$). Solving government *i*´s maximization problem under these conditions, and defining t_i^L and t_i^{PF} to be the capital taxes implemented by the double leader and the pure follower, respectively, in the Stackelberg equilibrium, produces the following results (see the Appendix);

Proposition 7: *Consider a federal system made up of two state governments and a federal government where state government j acts as a pure follower while state government i exercises double leadership. The marginal cost of public funds facing government i under double leadership will in this scenario be given by*

$$MCPF_i = \frac{\phi_i'' + \phi_j''}{\phi_j''} + \frac{1}{MCPF_j}\frac{\phi_i''}{u_j''} > 2 \tag{18}$$

Under double leadership, government i in this scenario implements a capital tax which eliminates the tax distortion within the fiscal union, i.e $t_i^L = t_j^{PF}$ *and* $\Delta = 0$, *where*

$$t_i^L = t_j^{PF} = -\left(\bar{k}_j - k_j(0)\right) f_i''\left(k_i(0)\right) \tag{19}$$

Observe first that the expression for the $MCPF_i$ in equation (18) is identical to, and can therefore be given the same interpretation as, the corresponding expression for the $MCPF_i$ in equation (16). This means that the under-taxation inefficiency is *qualitatively* equivalent with that outlined in Subsection 4.2. However, since the marginal cost of public funds in the other state appears in equations (16) and (18) and since $MCPF_j = 1$ if the other local government is a pure follower while $MCPF_j = 2$ if the other local government exercises decentralized leadership, there may very well be a substantial *quantitative* difference between the level of $MCPF_i$ in the two scenarios.

The key result in Proposition 7 is that government *i* implements a capital tax policy which eliminates the tax distortion within the fiscal union when the other state government is a pure follower. To explain this result, recall first from the discussion in Subsection 4.2 that both the pecuniary motive and the decentralized leadership motive for capital taxation vanishes when government i can exercise double leadership. This result also holds here. Recall also from Subsection 4.2 that when the other state government can exercise decentralized leadership vis-a-vis the federal government, then the capital allocation and the aggregate output are fixed at the levels implied by government j´s reaction function. However, the latter restriction is not present when the other government acts as a pure follower. The latter observation will have implications for government i´s capital tax policy. To see how, observe that since the lump-sum tax in state j satisfies the first-order condition $\phi_j' = u_j'$ while the federal transfer system satisfies the efficiency condition $\phi_i' = \phi_j'$, it follows that government i (when it exercises double leadership) recognizes the following chain of equalities; $\phi'(g_i) = \phi'\left(g_j\right) = u'\left(c_j\right)$. These equalities imply that the larger g_j and c_j are, the larger will be g_i. Effectively, this means that g_i will become an increasing function of the national income in the other state j. The latter observation provides government i with an incentive to implement a policy which maximizes the joint (i.e. aggregate) aggregate output within the fiscal union. This is accomplished by eliminating the tax distortion within the fiscal union and implies a tax policy where t_i^L is set equal to t_j^F so that $\Delta = t_i^L - t_j^F = 0$.

Since government i implements a capital tax that features $t_i^L = t_j^F$, there are two possible Stackelberg equilibria. When the pure follower is a net exporter of capital, it will implement a negative capital tax which the double leader replicates. This is illustrated in Figure 2 where the Stackelberg equilibrium is point C where both member states implement negative capital taxes. Note, however, that this low-tax equilibrium is not a "race-to-the-bottom" result because government i´s choice to implement a low capital tax when it exercises double leadership is not motivated by an incentive to attract domestic capital.

If the pure follower instead is a net importer of capital and therefore implements a positive capital tax, then the resulting Stackelberg equilibrium features positive capital taxes. These results can be summarized as follows;

Proposition 8: *Consider a federal system made up of two state governments and a federal government.*

- *(i)* *If the pure follower is the net exporter of capital then the Stackelberg equilibrium with double leadership features a low-tax equilibrium with negative capital taxes.*
- *(ii)* *If the pure follower is the net importer of capital then the Stackelberg equilibrium with double leadership features a high-tax equilibrium with positive capital taxes.*

6. Concluding Discussion

This paper presents a model of a two-layer fiscal union where the federal government uses lump-sum transfers to redistribute resources between two local jurisdictions and where each local government uses a capital tax and a lump-sum tax to finance the provision of a local public good. The federal government acts as a pure follower vis-a-vis the state governments but a state government may either act as a pure follower, exercise horizontal leadership and/or exercise decentralized leadership. The main novelty and contribution of the paper is to analyze the outcome when a state government simultaneously exercises horizontal and decentralized leadership. Another contribution is to extend the previous work on pure decentralized leadership in a fiscal union by allowing the member states to be heterogenous in terms of the endowments of capital and/or the production functions.

We started by analyzing the outcome when the local governments treat the choices made by the federal government as exogenous. In this context, we compared the Nash equilibrium, where each local government acts as pure follower and treats the other government´s tax policy as exogenously given, with the Stackelberg equilibrium where one of the local governments recognizes the other local government´s reaction function. The main result is that the tax distortion is smaller, and the welfare is larger, in the Nash equilibrium compared with the Stackelberg equilibrium with horizontal leadership. In the text, it was shown that this result is a consequence of the federal redistribution policy; without federal redistribution the welfare comparison would be inconclusive unless additional assumptions are made.

The second part of the paper concerned a fiscal union where both local governments are able to exercise decentralized leadership vis-a-vis the federal government. Here it was shown that each local government´s reaction function w.r.t. the capital tax has a slope equal to one in the capital tax space. This means that it is not possible to retrieve a unique Nash equilibrium w.r.t. the capital taxes; either the two reaction functions over-lap in which case there is an infinite number of capital tax combinations associated with a Nash equilibrium, or the two reaction functions are parallell to each other in which case no Nash equilibrium exists. When the two reaction functions over-lap, it was shown that there exists a unique Nash equilibrium w.r.t. the capital allocation between the two states either if the two states are identical so that the resulting Nash equilibrium is symmetric, or if the production functions in the two member states are quadratic and if the two production functions have identical second derivatives.

In the context of a fiscal union where both state govenrments exercise decentralized leadership vis-a-vis the federal government, we also characterized the outcome when one of the local governments is able to exercise horizontal leadership vis-a-vis the other local government. When a local government is able to exercise this type of double leadership, it was shown that the capital tax becomes redundant as a policy instrument for the double leader. The reason is that when the double leader takes into account the responses by the federal government and the other local government, then the double leader recognizes that it cannot influence the allocation of capital within the fiscal union. The explanation is that the double leader observes that there is a unique capital allocation associated with the other state government´s reaction function w.r.t. the capital tax. Another key result is that double leadership exacerbates the undertaxation inefficiency that earlier research has shown exists in a fiscal union with decentralized leadership.

Finally, we considered a fiscal union where one state government is able to exercise decentralized leadership while the other state government acts as a pure follower. In this context, it was shown that the double leader implements a capital tax which is equal to the capital tax implemented by the follower. This tax policy is economically efficient in the sense that it eliminates the tax distortion within the fiscal union and two equilibria are possible; a low-tax equilibrium where the double leader is a net importer of capital and a high-tax equilibrium where the double leader is a net exporter of capital.

Appendix

Derivation of Equation (7)

Differentiate equation (6) w.r.t. t_i. Then use that $r = f_i' - t_i$ implies $\partial r/\partial t_i = f_i''\partial k_i/\partial t_i - 1$ and that the first-order condition w.r.t. T_i implies $u_i' = \phi_i'$. This produces

$$\frac{\partial W_i}{\partial t_i} = u_i'\left(\left(\bar{k}_i - k_i\right)\frac{\partial r}{\partial t_i} + t_i\frac{\partial k_i}{\partial t_i}\right) = 0 \tag{A.1}$$

Solving for t_i in (A.1) produces the first equality in equation (7).

The Nash Equilibrium without Decentralized Leadership

From (8) we know that the slope of state i´s reaction function $t_i(t_j)$ is situated in the interval $0 < t_i'(t_j) < 1$. Analogously, equation (8) implies that the slope of state j´s reaction function $t_j(t_i)$ is situated in the interval $0 < t_j'(t_i) < 1$ which means that its slope in (t_j, t_i) - space is larger than one.

From (7), it follows that the intercept of state i´s reaction function at the horizontal axis (point D in Figure 1 where $t_i = 0$) is implicitly determined by the equation $0 = (\bar{k}_i - k_i)f_j''$. Since $f_j'' < 0$, it follows that $\bar{k}_i = k_i$ holds at this point which means that $t_i(t_j)$ also intersects with the ZCE locus at this point. Let us now ask whether the other state´s reaction function $t_j(t_i)$ also could go through point D? The answer is no because equation (7) implies that government *j* sets t_j according to the rule $t_j = (\bar{k}_j - k_j)f_i''$. Since $\bar{k}_j = k_j(0 - t_j^D)$ holds at point D, where t_j^D is state *j*´s capital tax at point D, it follows that the optimal tax rule would imply that $t_j = 0$ which is a contradiction ($t_j^D < 0$ at point D). Rather, from (A.1) (since we now consider government *j* we replace the index *i* with the index *j*, and vice versa, in A.1) it follows that $\partial W_j/\partial t_j > 0$ if government *j* implements the tax $t_j^D < 0$. Hence, government j´s optimal capital tax when $t_i = 0$ must be larger than t_j^D. Note also that if government *j* would set $t_j = 0$ when $t_i = 0$, then $\partial W_j/\partial t_j < 0$. Together, these arguments imply that the intercept of state *j*´s reaction function with the horizontal axis is a point such as C in Figure 1 where $t_j^D < t_j^C < 0$ and where t_j^C is state j´s capital tax at point C.

Using an analogous argument as in the preceding paragraph, it follows that the intercept of state j´s reaction function with the vertical axis (point A in Figure 1 where $t_j = 0$) is implicitly determined by the equation $0 = (\bar{k}_j - k_j) f_i''$. Since $f_i'' < 0$, it follows that $\bar{k}_j = k_j$ holds at this point which means that $t_j(t_i)$ also intersects with the ZCE locus at this point. Observe that the other state´s reaction function $t_i(t_j)$ cannot pass through point A because equation (7) implies that government *i* sets t_i according to $t_i = (\bar{k}_i - k_i) f_j''$. Since $\bar{k}_i = k_i(t_i^A - 0)$ holds at point A, where t_i^A is state i´s capital tax at point A, the optimal tax rule would imply that $t_i = 0$ which is a contradiction ($t_j^A > 0$ at point A). Rather, from (A.1) it follows that $\partial W_i/\partial t_i < 0$ if government *i* implements the tax $t_i^A > 0$. Hence, government i´s optimal capital tax when $t_j = 0$ must be smaller than t_j^A. Note also that if government *i* would set $t_i = 0$ when $t_j = 0$, then $\partial W_i/\partial t_i > 0$. Together, these arguments imply that the intercept of state i´s reaction function with the vertical axis is a point such as B in Figure 1 where $0 < t_i^B < t_i^A$ and where t_i^B is state i´s capital tax at point B.

Derivation of Equation (9)

Differentiate $W_i\left(t_i, T_i, t_j(t_i)\right)$ w.r.t. t_i. Next, use that $r = f_i' - t_i$ implies $\partial r/\partial t_j = f_i'' \partial k_i/\partial t_j$, together with $\partial r/\partial t_i = f_i'' \partial k_i/\partial t_i - 1$ and $u_i' = \phi_i'$, to rewrite the resulting expression to read

$$\frac{\partial W_i}{\partial t_i} = u_i'\left(\left(\bar{k}_i - k_i\right)\frac{\partial r}{\partial t_i} + t_i\frac{\partial k_i}{\partial t_i}\right) + u_i'\left(\left(\bar{k}_i - k_i\right)\frac{\partial r}{\partial t_j} + t_i\frac{\partial k_i}{\partial t_j}\right)\frac{\partial t_j}{\partial t_i} = 0 \tag{A.2}$$

Solving for t_i produces

$$t_i = -\left(\bar{k}_i - k_i\right)\frac{\left(\frac{\partial r}{\partial t_i} + \frac{\partial r}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right)}{\left(\frac{\partial k_i}{\partial t_i} + \frac{\partial k_i}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right)} \tag{A.3}$$

Using (4) and (8) in (A.3) produces equation (9).

Government i´s Maximization Problem under Pure Decentralized Leadership

When government *i* exercises pure decentralized leadership, the first-order conditions become

$$\frac{\partial W_i}{\partial T_i} = \phi_i'\left(1 - \frac{\partial R_i}{\partial T_i}\right) - u_i' = 0 \tag{A.4}$$

$$\frac{\partial W_i}{\partial t_i} = \phi_i'\left(k_i + t_i\frac{\partial k_i}{\partial t_i} - \frac{\partial R_i}{\partial t_i}\right) + u_i'\left(\bar{k}_i\frac{\partial r}{\partial t_i} - f_i'' k_i\frac{\partial k_i}{\partial t_i}\right) = 0 \tag{A.5}$$

Substituting $\partial R_i/\partial T_i = 2$ into (A.4) and then using the definition $MCPF_i = \phi_i'/u_i'$ produces equation (12). Next, multiply (A.4) by k_i, subtract the resulting expression from (A.5), use that $\partial r/\partial t_i = f_i'' \partial k_i/\partial t_i - 1$ and divide by u_i'. Then add and subtract $t_i\, \partial k_i/\partial t_i$. This produces

$$0 = t_i\frac{\partial k_i}{\partial t_i} + (MCPF_i - 1)t_i\frac{\partial k_i}{\partial t_i} - MCPF_i\left(\frac{\partial R_i}{\partial t_i} - k_i\frac{\partial R_i}{\partial T_i}\right) + \left(\bar{k}_i - k_i\right)\frac{\partial r}{\partial t_i} \tag{A.6}$$

where

$$\frac{\partial R_i}{\partial t_i} - k_i\frac{\partial R_i}{\partial T_i} = \frac{\phi_i'' t_i\frac{\partial k_i}{\partial t_i} - \phi_j'' t_j\frac{\partial k_j}{\partial t_i}}{\phi_i'' + \phi_j''} \tag{A.7}$$

Substitute equation (12) and equation (A.7) into (A.6). Solving for t_i in the resulting expression gives

$$t_i = -\left(\bar{k}_i - k_i\right)\frac{\partial r/\partial t_i}{\partial k_i/\partial t_i} - t_j\frac{\partial k_j/\partial t_i}{\partial k_i/\partial t_i} \tag{A.8}$$

Using the comparative static properties in (4) on the first term on the RHS produces equation (13) in the text.

Proof of Part (b) in Proposition 3

When the production functions are of the form $f_i(k_i) = a_i k_i - \frac{1}{2} b k_i^2$ and $f_j(k_j) = a_j k_j - \frac{1}{2} b k_j^2$, the functions $\Omega_i(k_i)$ and $\Omega_i(k_i)$ become

$$\Omega_i(k_i) = a_i - a_j + 2b(\bar{k}_i - k_i) + b(\bar{k}_j - k_i) = 0 \tag{A.9}$$
$$\Omega_j(k_j) = a_j - a_i + 2b(\bar{k}_j - k_j) + b(\bar{k}_i - k_j) = 0 \tag{A.10}$$

Using (from (3)) that $(\bar{k}_j - k_j) = -(\bar{k}_i - k_i)$ and that $(\bar{k}_i - k_j) = -(\bar{k}_j - k_i)$, respectively, in $\Omega_j(k_j) = 0$, and multiplying the resulting expression by -1, reproduces $\Omega_i(k_i) = 0$. Hence, $\Omega_i(k_i) = 0$ and $\Omega_j(k_j) = 0$ are identical and therefore imply identical capital allocations.

Government i´s Maximization Problem under Double Leadership vs a Decentralized Leader

When government *i* exercises double leadership, then it recognizes the other state government´s reaction function $t_j(t_i)$ together with the fact that R_i is determined by $0 = \phi'(g_j) - \phi'(g_i)$ and the fact that government *j*´s first-order condition for T_j is determined by (see (A.4)) $0 = 0.5\phi'(g_j) - u_j'(c_j)$ where $0.5 = 1/MCPF_j$ in government *j*´s first-order condition for T_j, and where

$$c_j = f_j\left(k_j\left(t_i - t_j(t_i)\right)\right) - f_j'\left(k_j\left(t_i - t_j(t_i)\right)\right) k_j\left(t_i - t_j(t_i)\right) + r\left(t_i, t_j(t_i)\right)\bar{k}_j - T_j \tag{A.11}$$
$$g_i = t_i k_i\left(t_i - t_j(t_i)\right) + T_i - R_i \tag{A.12}$$
$$g_j = t_j(t_i) k_j\left(t_i - t_j(t_i)\right) + T_j + R_i \tag{A.13}$$

We have also used that $R_j = -R_i$. To solve government *i*´s maximization problem when it exercises double leadership, we need to evaluate the effects of T_i and t_i on R_i. Therefore we differentiate $\phi'(g_j) - \phi'(g_i)$ and $0.5\phi'(g_j) - u_j'(c_j)$ w.r.t. T_i and t_i, respectively. This produces the following two equation systems (in matrix form)

$$\begin{bmatrix} (\phi_i'' + \phi_j'') & \phi_j'' \\ 0.5\phi_j'' & (0.5\phi_j'' + u_j'') \end{bmatrix} \cdot \begin{bmatrix} \partial R_i/\partial T_i \\ \partial T_j/\partial T_i \end{bmatrix} = \begin{bmatrix} \phi_i'' \\ 0 \end{bmatrix} \tag{A.14}$$
$$\begin{bmatrix} (\phi_i'' + \phi_j'') & \phi_j'' \\ 0.5\phi_j'' & (0.5\phi_j'' + u_j'') \end{bmatrix} \cdot \begin{bmatrix} \partial R_i/\partial t_i \\ \partial T_j/\partial t_i \end{bmatrix} = \begin{bmatrix} \phi_i'' k_i - \phi_j'' k_j \\ -u_j''\bar{k}_j - 0.5\phi_j'' k_j \end{bmatrix} \tag{A.15}$$

where we have used that the properties of the capital demand functions, the interest rate function and the fact that $t_j'(t_i) = 1$ imply $\partial c_j/\partial t_i = -\bar{k}_j$, $\partial c_j/\partial T_i = 0$, $\partial g_i/\partial t_i = k_i$, $\partial g_i/\partial T_i = 1$, $\partial g_j/\partial t_i = k_j$ and $\partial g_j/\partial T_i = 0$. Solving for $\partial R_i/\partial T_i$ and $\partial R_i/\partial t_i$ gives

$$\frac{\partial R_i}{\partial T_i} = \frac{\phi_i''(0.5\phi_j'' + u_j'')}{u_j''\phi_j'' + \phi_i''(0.5\phi_j'' + u_j'')}, \qquad \frac{\partial R_i}{\partial t_i} = \frac{\phi_i''(0.5\phi_j'' + u_j'')k_i + (\bar{k}_j - k_j)u_j''\phi_j''}{u_j''\phi_j'' + \phi_i''(0.5\phi_j'' + u_j'')} \tag{A.16}$$

Let us now return to government *i*´s maximization problem. The first-order conditions become

$$\frac{\partial W_i}{\partial T_i} = \phi_i'\left(1 - \frac{\partial R_i}{\partial T_i}\right) - u_i' = 0 \tag{A.17}$$
$$\frac{\partial W_i}{\partial t_i} = \phi_i'\left(k_i + t_i\left(\frac{\partial k_i}{\partial t_i} + \frac{\partial k_i}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) - \frac{\partial R_i}{\partial t_i}\right) + u_i'\bar{k}_i\left(\frac{\partial r}{\partial t_i} + \frac{\partial r}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) - u_i' k_i f_i''\left(\frac{\partial k_i}{\partial t_i} + \frac{\partial k_i}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) = 0 \tag{A.18}$$

First, substitute the expression for $\partial R_i/\partial T_i$ in (A.16) into (A.17) and then use that $MCPF_i = \phi_i'/u_i'$. Solving for $MCPF_i$ produces

$$MCPF_i = \frac{u_j''\phi_j'' + \phi_i''(0.5\phi_j'' + u_j'')}{u_j''\phi_j''} = 2 + 0.5\frac{\phi_i''}{u_j''} \tag{A.18}$$

Using that $0.5 = 1/MCPF_j$ in (A.18) produces equation (16) in the text. To evaluate the first-order condition for the capital tax (equation A.18), let us use the adding up conditions in (14) and (17), substitute the expression for $\partial R_i/\partial t_i$ in (A.16) into (A.18), divide by u_i' and use that $MCPF_i = \phi_i'/u_i'$. We obtain

$$0 = MCPF_i\left[k_i - \frac{\phi_i''\left(0.5\phi_j''+u_j''\right)k_i+(\bar{k}_j-k_j)u_j''\phi_j''}{u_j''\phi_j''+\phi_i''\left(0.5\phi_j''+u_j''\right)}\right] - \bar{k}_i \tag{A.19}$$

Write the expression inside square brackets on a common denominator, substitute in the (middle) expression for $MCPF_i$ in (A.18) into (A.19), simplify and multiply all terms by $-u_j''\phi_j''$. This produces

$$0 = \left(\bar{k}_i + \bar{k}_j - k_i - k_j\right)u_j''\phi_j'' \tag{A.20}$$

Since the expression inside brackets is the capital market equilibrium condition (see equation 3), equation (A.20) is satisfied for any t_i. This verifies Proposition 5.

Government i´s Maximization Problem under Double Leadership vs a Pure Follower

Government i´s maximization problem is the same as above except that the slope of $t_j(t_i)$ is now less than one and T_j is now determined by $0 = 1\cdot\phi'(g_j) - u_j'(c_j)$, where we observe that $1 = 1/MCPF_j$ (because $MCPF_j = 1$ when government j acts as a pure follower) in government j´s first-order condition for T_j. As above, R_i is determined by $0 = \phi'(g_j) - \phi'(g_i)$. Differentiating the latter first-order conditions w.r.t. T_i and t_i produces the following two equation systems

$$\begin{bmatrix} (\phi_i'' + \phi_j'') & \phi_j'' \\ 1\cdot\phi_j'' & (1\cdot\phi_j'' + u_j'') \end{bmatrix} \cdot \begin{bmatrix} \partial R_i/\partial T_i \\ \partial T_j/\partial T_i \end{bmatrix} = \begin{bmatrix} \phi_i'' \\ 0 \end{bmatrix} \tag{A.21}$$

$$\begin{bmatrix} (\phi_i'' + \phi_j'') & \phi_j'' \\ 1\cdot\phi_j'' & (1\cdot\phi_j'' + u_j'') \end{bmatrix} \cdot \begin{bmatrix} \partial R_i/\partial t_i \\ \partial T_j/\partial t_i \end{bmatrix} = \begin{bmatrix} \phi_i''\frac{\partial g_i}{\partial t_i} - \phi_j''\frac{\partial g_j}{\partial t_i} \\ u_j''\frac{\partial c_j}{\partial t_i} - 1\cdot\phi_j''\frac{\partial g_j}{\partial t_i} \end{bmatrix} \tag{A.22}$$

where

$$\frac{\partial g_i}{\partial t_i} = k_i + t_i\left(\frac{\partial k_i}{\partial t_i} + \frac{\partial k_i}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) \tag{A.23}$$

$$\frac{\partial g_j}{\partial t_i} = k_j\frac{\partial t_j}{\partial t_i} + t_j\left(\frac{\partial k_j}{\partial t_i} + \frac{\partial k_j}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) \tag{A.24}$$

$$\frac{\partial c_j}{\partial t_i} = -f_j''k_j\left(\frac{\partial k_j}{\partial t_i} + \frac{\partial k_j}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) + \bar{k}_j\left(\frac{\partial r}{\partial t_i} + \frac{\partial r}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) \tag{A.25}$$

Solving for $\partial R_i/\partial T_i$ and $\partial R_i/\partial t_i$ gives

$$\frac{\partial R_i}{\partial T_i} = \frac{\phi_i''\left(1\cdot\phi_j''+u_j''\right)}{u_j''\phi_j''+\phi_i''\left(1\cdot\phi_j''+u_j''\right)}, \qquad \frac{\partial R_i}{\partial t_i} = \frac{\phi_i''\left(1\cdot\phi_j''+u_j''\right)\frac{\partial g_i}{\partial t_i} - \phi_j''u_j''\left(\frac{\partial g_j}{\partial t_i}+\frac{\partial c_j}{\partial t_i}\right)}{u_j''\phi_j''+\phi_i''\left(1\cdot\phi_j''+u_j''\right)} \tag{A.26}$$

Let us now return to government i´s maximization problem. The first-order conditions become

$$\frac{\partial W_i}{\partial T_i} = \phi_i'\left(1 - \frac{\partial R_i}{\partial T_i}\right) - u_i' = 0 \tag{A.27}$$

$$\frac{\partial W_i}{\partial t_i} = \phi_i'\left(\frac{\partial g_i}{\partial t_i} - \frac{\partial R_i}{\partial t_i}\right) + u_i'\frac{\partial c_i}{\partial t_i} = 0 \tag{A.28}$$

where

$$\frac{\partial c_i}{\partial t_i} = -f_i''k_i\left(\frac{\partial k_i}{\partial t_i} + \frac{\partial k_i}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) + \bar{k}_i\left(\frac{\partial r}{\partial t_i} + \frac{\partial r}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) \tag{A.29}$$

and where $\partial g_i/\partial t_i$ is defined in (A.23) (and where we note that the effects via $\partial R_i/\partial t_i$ are not incorporated in A.23). First, substitute the expression for $\partial R_i/\partial T_i$ in (A.26) into (A.27) and then use that $MCPF_i = \phi_i'/u_i'$. Solving for $MCPF_i$ produces

$$MCPF_i = \frac{u_j''\phi_j''+\phi_i''\left(1\cdot\phi_j''+u_j''\right)}{u_j''\phi_j''} = 2 + 1\cdot\frac{\phi_i''}{u_j''} \tag{A.30}$$

Using that $1 = 1/MCPF_j$ in (A.30) reproduces equation (18) in the text. Turning to (A.28), let us divide by u_i' and use that $MCPF_i = \phi_i'/u_i'$. This produces

$$0 = MCPF_i\left(\frac{\partial g_i}{\partial t_i} - \frac{\partial R_i}{\partial t_i}\right) + \frac{\partial c_i}{\partial t_i} \tag{A.31}$$

Next, substitute the middle expression in (A.30), and the expression for $\partial R_i/\partial t_i$, into (A.31). Simplifying the resulting expression produces

$$0 = \frac{\partial g_i}{\partial t_i} + \frac{\partial g_j}{\partial t_i} + \frac{\partial c_j}{\partial t_i} + \frac{\partial c_i}{\partial t_i} \tag{A.32}$$

Substituting in equation (A.23) – (A.25) and (A.29) into (A.32) gives

$$0 = k_i + k_j\frac{\partial t_j}{\partial t_i} + (t_i - f_i''k_i)\left(\frac{\partial k_i}{\partial t_i} + \frac{\partial k_i}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) + (t_j - f_j''k_j)\left(\frac{\partial k_j}{\partial t_i} + \frac{\partial k_j}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) + (\bar{k}_i + \bar{k}_j)\left(\frac{\partial r}{\partial t_i} + \frac{\partial r}{\partial t_j}\frac{\partial t_j}{\partial t_i}\right) \tag{A.33}$$

Use that the comparative static properties in (4) imply

$$\frac{\partial k_i}{\partial t_i} + \frac{\partial k_i}{\partial t_j}\frac{\partial t_j}{\partial t_i} = \frac{1-\frac{\partial t_j}{\partial t_i}}{f_i''+f_j''}, \qquad \frac{\partial k_j}{\partial t_i} + \frac{\partial k_j}{\partial t_j}\frac{\partial t_j}{\partial t_i} = -\frac{1-\frac{\partial t_j}{\partial t_i}}{f_i''+f_j''}, \qquad \frac{\partial r}{\partial t_i} + \frac{\partial r}{\partial t_j}\frac{\partial t_j}{\partial t_i} = -\frac{f_i''\frac{\partial t_j}{\partial t_i}+f_j''}{f_i''+f_j''} \tag{A.34}$$

Substituting the expressions in (A.34) into (A.33) and collecting terms produces

$$0 = (t_i - t_j)\left(\frac{1-\frac{\partial t_j}{\partial t_i}}{f_i''+f_j''}\right) - (\bar{k}_i + \bar{k}_j - k_i - k_j)\left(\frac{f_i''\frac{\partial t_j}{\partial t_i}+f_j''}{f_i''+f_j''}\right) \tag{A.35}$$

The second term on the RHS is zero (follows from equation 3). Since $1 - \partial t_j/\partial t_i > 0$ and $f_i'' + f_j'' < 0$, it follows that (A.35) can only be satisfied if t_i is set equal to t_j which, in turn, implies $\Delta = 0$. Equations (2) and (3) then implicitly determine the Stackelberg equilibrium capital stock, k_i^L, in state *i* as $f_i'(k_i^L) = f_j'(\bar{k}_i + \bar{k}_j - k_i^L)$. Using that equation (7) determines the capital tax for the pure follower, it follows that the Stackelberg equilibrium capital taxes are determined by $t_i^L = t_j^F = (\bar{k}_j - k_j^F)f_i''(k_i^L)$. Hence, if state *i* is a net importer of capital (i.e. $\bar{k}_i - k_i(\Delta = 0) < 0$), it follows that the equilibrium taxes are negative, as illustrated in Figure 2. If state *i* instead is a net exporter of capital (i.e. $\bar{k}_i - k_i(\Delta = 0) > 0$), then the equilibrium taxes are positive, as illustrated in Figure 3.

References

Altshuler, R. And Goodspeed, T.J. (2015), "Follow the Leader ? Evidence on European and US Tax Competition", *Public Finance Review* 43 (4), 485-504.

Aronsson, T., Jonsson, T. And Sjögren, T. (2006), "Environmental Policy and Optimal Taxation in a Decentralized Economic Federation", *FinanzArchiv* 62 (3), 437-454.

Bucovetsky, S. (2009) "An Index of Capital Tax Competition", *International Tax Public Finance* 16, 727-752.

Bucovetsky, S. and Wilson, J.D. (1991), "Tax Competition with Two Tax Instruments", *Regional Science and Urban Economics* 21, 333-350.

Caplan, A.J. and Silva, E.C.D. (1999), "Federal Acid Rain Games", *Journal of Urban Economics* 46, 25–52.

Caplan, A.J., Cornes, R.C. and Silva, E.C.D. (2000), "Pure Public Goods and Income Redistribution in a Federation with Decentralized Leadership and Imperfect Labor Mobility", *Journal of Public Economics* 77, 265–284.

DePater, J.A. and Myers, G.M. (1994), "Strategic Capital Tax Competition. A Pecuniary Externality and a Corrective Device", *Journal of Urban Economics* 36, 66-78.

Eichner, T. (2014), “Endogenizing Leadership and Tax Competition: Externalities and Public Good Provision“, *Regional Science and Urban Economics* 46, 18-26.

Gordon, R. (1992), “Can Capital Income Taxes Survive in Open Economies ?“, *Journal of Finance* 47, 1159-1180.

Hamilton, J.H. and Slutsky, S.M. (1990) “Endogenous Timing in Duopoly Games: Stackelberg or Cournot Equilibria“, *Games and Economic Behavior* 2 (1), 29–46.

Hoyt, W.H. (2013), “Tax Policy Coordination, Vertical Fiscal Externalities, and Optimal Taxation in a System of Hierarchical Governments“, *Journal of Urban Economics* 50, 491-516.

Itaya, J., Okamura, M. and Yamagucchi, C. (2008), “Are Regional Asymmetries Detrimental to Tax Coordination in a Repeated Game Setting?“, *Journal of Public Economics* 92, 2403-2411.

Keen, M. and Marchand, M. (1997), “Fiscal Competition and the Pattern of Public Spending“, *Journal of Public Economics* 66 (1), 33-53.

Keen M. And Konrad, K. (2013), “The Theory of International Tax Competition and Tax Coordination.“ Chetty, A.J., Feldstein M. And Saez E. (Eds), Handbook of Public Economics 5, 257-328.

Kempf, H. and Rota-Graziosi, G. (2010), “Endogenizing Leadership in Tax Competition“, *Journal of Public Economics* 94, 768-776.

Konrad, K. and Schjelderup, G. (1999), “Fortress Building in Global Tax Competition“, *Journal of Urban Economics* 46, 156-167.

Köthenbürger, M. (2004), “Tax Competition in a Fiscal Union with Decentralized Leadership“, *Journal of Urban Economics* 55, 498-513.

Ogawa, H. (2013), “Further Analysis on Leadership in Tax Competition: The Role of Capital Ownership“, *International Tax and Public Finance* 20, 474-484.

Peralta, S. and van Ypersele, T. (2005) “Factor Endowments and Welfare Levels in an Asymmetric Tax Competition Game“, *Journal of Urban Economics* 57 (2), 258–274.

Schelling, T. (1960), “The Strategy of Conflict.“ Harvard University Press, Cambridge, MA.

Silva, E.C.D. and Caplan, A.J. (1997), “Transboundary Pollution Control in Federal Systems“, *Journal of Environmental Economics and Management* 34, 173–186.

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

Vives, X. (1999), “Oligopoly Pricing. Old Ideas and New Tools.” The MIT Press, Cambridge, Massachusetts.

Wildasin, D.E. (1989), “Interjurisdictional Capital Mobility: Fiscal Externality and Corrective Subsidy“, *Journal of Urban Economics* 25, 193–212.

Wilson, J.D. (1999), “Theories of Tax Competition“, *National Tax Journal* 52 (2), 269–304.