Evolutionary equilibria in capital tax competition with imitative learning

Hiroyuki Sano

No.0603

OUC Economic Research Papers

Department of Economics

Otaru University of Commerce

Evolutionary equilibria in capital tax competition with imitative learning

Hiroyuki Sano*

Department of Economics, Otaru University of Commerce,

3-5-21, Midori, Otaru, 047-8501 Japan

September 30, 2005

Abstract

This paper presents an alternative model where each decision maker takes an economy-wide net return to capital parametrically even in tax competition among finite nonatomistic jurisdictions. To this end, we examine a long-run equilibrium of a tax competition game adopting an evolutionary game-theoretic approach. We find that a unique evolutionarily stable strategy coincides with the strategy consisting of a unique stochastically stable state in the imitative dynamics of tax competition. Moreover, in the evolutionary equilibria, we obtain the same result as in the "purely competitive" equilibrium even with finite nonatomistic jurisdictions.

***JEL Classification*:** C72, H77.

***Keywords*:** Tax competition, Imitative learning, Evolutionarily stable strategy, Stochastically stable state.

* *Tel & fax*: +81-134-27-5304. *E-mail address*: sano@res.otaru-uc.ac.jp

1 Introduction

The standard tax competition model constructed by Zodrow and Mieszkowski (1986) and Wilson(1986) assumes that a decision maker in each jurisdiction takes an economy-wide net return to capital parametrically. In other words, each jurisdiction is assumed to be very small in that a jurisdictional tax policy cannot affect the net return to capital and thus has no effect on residents' utilities in other jurisdictions. In the "purely competitive" equilibrium with atomistic jurisdictions, capital tax competition results in inefficiently low tax rates and local public service levels.

Wildasin(1988) studies capital tax competition in the case where each jurisdiction affects residents' utilities in other jurisdictions by changing the economy-wide net return to capital. With this strategic consideration, the size of each jurisdiction or equivalently the number of jurisdictions is important. Each jurisdiction is assumed to be large enough for a unilateral change in the tax rate to affect the amounts of capital employed in other jurisdictions positively or negatively. Hoyt(1991) shows that the capital tax rate of each jurisdiction in a Nash equilibrium decreases with the number of jurisdictions, and converges to the tax rate in the purely competitive equilibrium with small atomistic jurisdictions as the number of jurisdictions goes to infinity. The latter result implies that the purely competitive equilibrium is equivalent to the Nash equilibrium among atomistic jurisdictions.

In this paper, we are concerned with an alternative model where each decision maker takes the economy-wide net return to capital parametrically even in tax competition among finite nonatomistic jurisdictions. We examine the long-run equilibrium of a tax competition game adopting an evolutionary game-theoretic approach in place of the classic game-theoretic approach where the Nash equilibrium concept is commonly utilized.

In the standard tax competition model, it is assumed that each local government attempts to maximize the utility of a representative resident in its jurisdiction. The representative resident should be interpreted as a swing voter in either a direct or a representative democracy, such as a median voter in majority voting. When the swing voter who is an individual resident chooses tax policies maximizing his or her utility, he or she would not usually have enough of the information about a capital market that is needed to compute his or her own utility and levels of voters' utilities in other jurisdictions given any profile of tax rates. Even though the voter can acquire the information, he or she might not have the ability to compute levels of those utilities accurately. It would be inappropriate to describe the situations where the voters do not have enough information and/or computation ability using a Nash game among rational agents. Also, the Nash equilibrium concept is not plausible in that it requires common knowledge of rationality among voters across all jurisdictions. To deal with a tax competition game among boundedly rational voters, therefore, we apply a concept of equilibrium selection in an evolutionary setting rather than the Nash equilibrium concept.

Boundedly rational voters need to learn from past experiences to choose their tax policies. Learning behavior has two classes: one is learning from one's own jurisdiction's experiences only, and the other is learning from the federation-wide experiences. The former is called *introspective learning* and the latter is *imitative learning*. In this paper, we consider imitative learning behaviors in a discrete-time dynamic setting. More specifically, we assume that in the next period the boundedly rational voters imitate the currently most successful tax policy. Behavioral assumptions in many local public finance literature are closer to imitative learning behaviors than introspective learning. For example, in Tiebout's(1956) model and those of his successors, individual residents are assumed to move into the jurisdiction in which the highest level of utility can be attained. In Besley and Case(1995) and Wrede(2001), who studied yardstick competition among self-interested governments, the voters compare performances of other governments with their own government in order to determine whether to reelect the incumbents or not. Thus, it would be natural that we focus on imitative learning behaviors rather than introspective learning.[1]

There have been few studies in local public-finance literature that attempted to deal with a game among boundedly rational voters as long as we know. In contrast, in industrial organization there is much research on various types of oligopolistic competition among boundedly rational agents. Schaffer(1989) describes Cournot competition as a Darwinian model of economic natural selection and shows that the evolutionarily stable strategy (ESS) which survives the selection process is not any Cournot–Nash equilibrium strategy but yields an efficient level of output. Vega-Redondo(1997) models learning dynamics of Cournot competition explicitly. He also proves that the efficient level of output is stochastically stable in the perturbed imitative dynamics where oligopolists imitate the currently most successful output. Alós-Ferrer, Ania and Schenk-Hoppé(2000)

[1] If the voters are not able to observe other jurisdictions' tax rates and/or private and public service consumptions, then the presumption of introspective learning is more appropriate than imitative learning. Bergin and Bernhardt(2004) investigate both cases in their setting.

investigate Bertrand competition in their imitative dynamics following the model of Vega-Redondo(1997).

We consider the two concepts of equilibrium selection. The first is a finite population ESS defined in Schaffer(1988). When using the definition of a standard ESS, our analysis is confined to the case of an 'infinite' population with infinitesimally small agents. Hence, we use the definition of the finite population ESS because it is suitable for our analysis of capital tax competition with finite nonatomistic jurisdictions. Although the ESS is a central concept of equilibrium selection, it does not provide an explicit dynamic selection process. Therefore, it is not certain whether imitative learning behaviors are the main driving forces in the selection process in which an ESS survives. Secondly, taking this drawback into account, we investigate a stochastically stable state (SSS) adopted in Foster and Young(1990), Kandori, Mailath and Rob(1993) and Young(1993). The stochastically stable state is a state which belongs to the support of the limit invariant distribution of a perturbed Markov process. We formalize imitative dynamics of tax competition as a specific perturbed Markov process along the lines of the approach by Vega-Redondo(1997). We will find that a unique ESS in our tax competition game coincides with the strategy consisting of a unique SSS in the imitative dynamics of tax competition. Moreover, in both investigations, we will obtain the same result as in the purely competitive equilibrium even with finite nonatomistic jurisdictions.

The rest of this paper is organized as follows. Section 2 presents the basic model. In Section 3, we derive a Nash equilibrium tax rate and verify its existence. Section 4 deals with a finite population ESS. In Section 5, we formalize imitative dynamics of tax competition and find out a stochastically stable state, and Section 6 concludes.

2 The basic model

Consider a federal state with a set of jurisdictions $J = \{1, 2, \ldots, n\}$. Each of the jurisdictions is inhabited by the identical number of individuals who are immobile among jurisdictions. Total number of individuals in the federal state is assumed to be a fixed $N \geq n$. The total amount of capital stock in the economy, denoted by $\bar{K}$, is also constant. Competitive firms in jurisdiction $i \in J$ produce a homogeneous private good x combining capital that is perfectly mobile across jurisdictions with a fixed immobile factor of labor. Each resident is endowed with one unit of labor that is always employed. The production of the private good per capita is given by $f(k_i)$ where k_i denotes the per capita capital stock located within jurisdiction i. Furthermore, we assume that the production function is quadratic,

$$f(k_i) = \left(a - \frac{b}{2}k_i\right)k_i, \tag{1}$$

where a and b are positive parameters. Thus, we have $f'(k_i) = a - bk_i$ and $f''(k_i) = -b$.[2] To ensure that $f' > 0$, it is assumed that $a \geq bn\bar{k}$ where $\bar{k} \equiv \bar{K}/N$.

A source-based tax can be imposed on a unit of capital employed in jurisdiction i at rate τ_i. Perfect mobility of capital across jurisdictions results in a common net (or after-tax) return to capital, denoted by ρ. Therefore, profit-maximizing behaviors of competitive firms yield

$$f'(k_i) = a - bk_i = \rho + \tau_i \text{ for every } i \in J. \tag{2}$$

The capital market-clearing condition $\sum_{j=1}^{n} k_j = n\bar{k}$ together with (2), we obtain

$$\rho = a - b\bar{k} - \bar{\tau}, \tag{3}$$

$$k_i = \begin{cases} \bar{k} + \frac{\bar{\tau}-\tau_i}{b} & \text{if } \tau_i < \frac{nb\bar{k}}{n-1} + \bar{\tau}_{-i}, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where $\bar{\tau} \equiv \sum_{j=1}^{n} \tau_j/n$, i.e., an average rate of capital tax, and $\bar{\tau}_{-i}$ is an average tax rate among jurisdictions except for i. Hence, the net return to capital ρ at the equilibrium of the capital market depends on the average tax rate $\bar{\tau}$. Differentiating (3) with respect to τ_i for any $i \in J$ yields

$$\rho' \equiv \frac{d\rho}{d\tau_i} = -\frac{1}{n}. \tag{5}$$

Moreover, differentiating k_i with respect to τ_i and τ_j $(j \neq i)$, respectively, we obtain

$$\frac{\partial k_i}{\partial \tau_i} = -\frac{n-1}{nb} < 0, \tag{6a}$$

$$\frac{\partial k_i}{\partial \tau_j} = \frac{1}{nb} > 0, \tag{6b}$$

when $\tau_i < nb\bar{k}/(n-1) + \bar{\tau}_{-i}$. Eq. (6a) represents the degree of capital outflow from jurisdiction i by an infinitesimal increase in the tax rate of jurisdiction i. Conversely, capital

[2] Henceforth, a prime attached to a function with one variable means the derivative of the function.

inflow to jurisdiction i occurs when jurisdiction $j \neq i$ raises the tax rate infinitesimally. The degree of capital inflow can be seen in (6b).

Regional government $i \in J$ provides a local public service g_i that is assumed to be a publicly provided private good with complete rivalry.[3] Assume that producing one unit of the public service requires one unit of the private good as a numeraire good. To provide the local public service, regional government i can use only a source-based capital tax. Thus, using (4), the budget constraint per resident for government i is given by

$$g_i = \tau_i k_i = \tau_i \bar{k} + \frac{\tau_i \left(\bar{\tau} - \tau_i\right)}{b}, \tag{7}$$

if $\tau_i < nb\bar{k}/\left(n-1\right) + \bar{\tau}_{-i}$, otherwise $g_i = 0$. The rate of the source-based capital tax is determined through majority voting. Furthermore, government officials or politicians are assumed to be unable to acquire their ego rents.

It is assumed that each individual in the economy has an identical preference for private and local public-service consumption. The identical individual preference is described by the utility function $u\left(x_i, g_i\right)$ which is at least twice-continuously differentiable. Let the marginal rate of substitution be denoted by $mrs^i\left(x_i, g_i\right) \equiv u_g^i/u_x^i$. Furthermore, we need the following assumptions.

Assumption 1. For every i, $u_x^i \equiv \frac{\partial}{\partial x_i} u\left(x_i, g_i\right) > 0$, $u_g^i \equiv \frac{\partial}{\partial g_i} u\left(x_i, g_i\right) > 0$, $u_{xx}^i \equiv \frac{\partial^2}{\partial x_i^2} u\left(x_i, g_i\right) \leq 0$ and $u_{gg}^i \equiv \frac{\partial^2}{\partial g_i^2} u\left(x_i, g_i\right) \leq 0$.

Assumption 2. For every i and for all $x_i, g_i > 0$, $\frac{\partial}{\partial x_i} mrs^i\left(x_i, g_i\right) \geq 0$ and $\frac{\partial}{\partial g_i} mrs^i\left(x_i, g_i\right) \leq 0$.

Assumption 3. For every i and for all $x_i > 0$, $mrs^i\left(x_i, 0\right) > 1$.

Assumption 1 requires that the utility function is monotonically increasing for both goods and that neither marginal utility is increasing. Assumption 2 formally states that neither good is inferior. Assumption 3 implies that each individual is willing to forgo more than one unit of the private good to consume the first unit of the public service.[4]

[3] Although we can also assume a public service with more general properties, introducing nonrivalry of the local public service complicates the investigation on the effects of a change in the jurisdictions' size on tax competition.

[4] If we assume that $mrs^i\left(x_i, 0\right) \leq 1$, the tax rate in the efficient allocation is always zero since the marginal rate of transformation between the public services and the private goods is equal to one. This case is trivial for investigations of a tax competition game because a Nash equilibrium tax rate is also zero when $mrs^i\left(x_i, 0\right) \leq 1$.

All individuals are assumed to be initially endowed with the amount of capital equal to the average capital endowment $\bar{k} \equiv \bar{K}/N$. Using (3) and (4), the budget constraint for each individual in jurisdiction i is

$$x_i = f(k_i) - f'(k_i)\, k_i + \rho\bar{k} = f\left(\bar{k}\right) - \tau_i\bar{k} + \frac{(\bar{\tau} - \tau_i)^2}{2b}, \tag{8}$$

if $\tau_i < nb\bar{k}/(n-1) + \bar{\tau}_{-i}$, otherwise $x_i = \rho\bar{k}$. Differentiating (7) and (8) with respect to $\tau_i < nb\bar{k}/(n-1) + \bar{\tau}_{-i}$, respectively, we obtain

$$\frac{\partial g_i}{\partial \tau_i} = \bar{k} + \frac{\bar{\tau} - \tau_i}{b} - \frac{n-1}{n}\frac{\tau_i}{b} = k_i - \frac{n-1}{n}\frac{\tau_i}{b}, \tag{9}$$

$$\frac{\partial x_i}{\partial \tau_i} = -\left(\bar{k} + \frac{n-1}{n}\frac{\bar{\tau} - \tau_i}{b}\right) = -\frac{1}{n}\left[(n-1)\, k_i + \bar{k}\right] < 0. \tag{10}$$

We can observe from (9) that a unilateral increase in the tax rate at $\tau_i = \bar{\tau}$ raises the local public service if $\bar{k} - \frac{n-1}{n}\frac{\bar{\tau}}{b} > 0$. Eq.(10) shows that the private consumption always decreases with the tax rate.

Finally, both g_i and x_i depend on $\sum_{j \neq i} \tau_j$ as well as τ_i, thus the indirect utility for a resident in i is represented as $V\left(\tau_i, \sum_{j \neq i} \tau_j\right)$.

3 Symmetric Nash equilibrium

In this section, we characterize a symmetric Nash equilibrium among rational voters who have common knowledge of rationality and are able to compute their outcomes contingent on tax policies accurately with full information. Because all individuals are identical with respect to their preferences and endowments, any individual in each jurisdiction becomes a median voter who can force the regional government to choose his or her most preferred tax rate through majority voting.

Lemma 1 *Under assumptions 1–2 and the quadratic production function, the indirect utility function* $V\left(\tau_i, \sum_{j \neq i} \tau_j\right)$ *has a single peak with respect to* τ_i *given any* $\sum_{j \neq i} \tau_j$*.*

Proof. See Appendix. ■

Lemma 1 states that there is the only one capital tax rate chosen by jurisdiction i through majority voting given $\sum_{j \neq i} \tau_j$. Henceforth, the optimal tax rate for i denoted by

$\tau_i\left(\sum_{j\neq i}\tau_j\right)$. Bucovetsky (1991) showed that there exists a Nash equilibrium among two nonidentical jurisdictions when the production function is assumed to be quadratic. Similarly, given the quadratic production function defined above, we can obtain the existence and uniqueness of a symmetric Nash equilibrium among finite n identical jurisdictions.

Proposition 1 *Let assumptions 1–3 hold and the production functions be quadratic. Then, there exists a unique symmetric Nash equilibrium in the capital tax-competition model with finite identical jurisdictions.*

Proof. Suppose that the median voter in jurisdiction i chooses $\tau_i \in C$ so as to maximize $V\left(\tau_i,(n-1)\tau^{SNE}\right)$, where C is a continuum of tax rates on the interval $[0,\infty)$ in $\mathbb{R}_+$ and τ^{SNE} denotes the tax rate in a symmetric Nash equilibrium. The first-order condition for an interior solution of this maximization problem is, using (9) and (10),

$$\frac{u_g^i}{u_x^i}=\frac{k_i-\frac{1}{n}\left(k_i-\bar{k}\right)}{k_i-\frac{n-1}{n}\frac{\tau_i}{b}} \text{ for any } i\in J. \tag{11}$$

Thus, we have the following necessary condition for a symmetric Nash equilibrium:

$$mrs^i\left(x_i,g_i\right)=\frac{\bar{k}}{\bar{k}-\frac{n-1}{n}\frac{\tau^{SNE}}{b}} \text{ for any } i\in J. \tag{12}$$

The consumptions for both goods in the symmetric Nash equilibrium are, respectively,

$$x^{SNE}=f\left(\bar{k}\right)-\tau^{SNE}\bar{k} \text{ and } g^{SNE}=\tau^{SNE}\bar{k}. \tag{13}$$

Lemma 1 implies that for given any $\sum_{j\neq i}\tau_j$, there is at most one tax rate as the best response that corresponds to the bliss point of the median voter in jurisdiction i. Therefore, if τ^{SNE} satisfies the condition (12), then for every $i\in J$, $V\left(\tau^{SNE},(n-1)\tau^{SNE}\right)>V\left(\tau_i,(n-1)\tau^{SNE}\right)$ for any $\tau_i\in C\backslash\left\{\tau^{SNE}\right\}$.

Next, we verify that τ^{SNE} satisfying (12) uniquely exists. The right-hand side of (12) is monotonically increasing in τ^{SNE}, and moreover it is equal to 1 when $\tau^{SNE}=0$ and goes to infinity as τ^{SNE} approaches to $nb\bar{k}/(n-1)$. We claim that the left-hand side of (12), i.e., $mrs^i\left(x_i,g_i\right)$ is not increasing in τ^{SNE} owing to Assumption 2. Suppose that all jurisdictions carry out coordinated increases in τ^{SNE}. Then, differentiating x^{SNE} and

g^{SNE} in (13) with respect to τ^{SNE}, we obtain

$$\frac{\partial x^{SNE}}{\partial \tau^{SNE}} = -\bar{k} \text{ and } \frac{\partial g^{SNE}}{\partial \tau^{SNE}} = \bar{k}, \tag{14}$$

respectively. Thus, using (14), the effect of coordinated increases in τ^{SNE} on $mrs^i(x_i, g_i)$ is

$$\frac{\partial}{\partial \tau^{SNE}} mrs^i(x_i, g_i) = \left(\frac{\partial mrs^i}{\partial g_i} - \frac{\partial mrs^i}{\partial x_i} \right) \bar{k} \leq 0, \tag{15}$$

since $\partial mrs^i / \partial g_i \leq 0$ and $\partial mrs^i / \partial x_i \geq 0$ by Assumption 2. When $\tau^{SNE} = 0$, we have $g_i = 0$ so that $mrs^i(x, 0) > 1$ from Assumption 3, while the right-hand side of (12) is equal to unity. As τ^{SNE} approaches $nb\bar{k}/(n-1)$, the right-hand side of (12) goes to infinity. Hence, we can obtain a unique τ^{SNE} lying strictly between 0 and $nb\bar{k}/(n-1)$. ■

We can now investigate the effect of an increase in n on the symmetric Nash equilibrium. That n increases might be regarded as fiscal decentralization because it means the increase in the number of jurisdictions that can independently determine some policies for taxation and public expenditures. Totally differentiating (12) yields

$$\frac{d\tau^{SNE}}{dn} = \tau^{SNE} \bar{k} \left(n\bar{k} - \frac{n-1}{b} \tau^{SNE} \right)^{-2} \left[b \frac{\partial mrs^i}{\partial \tau^{SNE}} - \frac{(n-1)\bar{k}}{n \left(\bar{k} - \frac{n-1}{n} \frac{\tau^{SNE}}{b} \right)^2} \right]^{-1} \tag{16}$$

Since $\partial mrs^i / \partial \tau^{SNE} \leq 0$ from (15) and $b > 0$, we have $d\tau^{SNE}/dn < 0$, that is, τ^{SNE} decreases as n increases. Intuitively, fiscal decentralization lowers the equilibrium tax rate because tax competition becomes more intense. The same property is derived in Hoyt (1991).

Some tax-competition literature (e.g., Wilson [1986] and Zodrow and Mieszkowski [1986]) considers a case of many small identical jurisdictions, i.e., a purely competitive model. Assuming that $\rho' = 0$, we can capture the purely competitive case. Let a symmetric Nash equilibrium in this case be denoted by τ^*. Then, using (9) and (10), τ^* is such that

$$mrs^i(x^*, g^*) = \frac{\bar{k}}{\bar{k} - \frac{\tau^*}{b}}. \tag{17}$$

This expression can be also attained as $n \to \infty$ in (12). Therefore, it is clear from the proof of Proposition 1 that the symmetric Nash equilibrium tax rate satisfying (17) is uniquely determined. Comparing (17) with (12), the tax rate in the purely competitive equilibrium is lower than in the Nash equilibrium with any finite number of jurisdictions.

4 Economic natural selection

This section investigates an evolutionarily stable strategy (ESS) in the capital tax competition game. The analysis of an ESS is generally carried out in the context where there is an infinite population in that the agents are infinitesimal relative to the size of the population. From the viewpoint of reality, however, it would not be appropriate to assume that each jurisdiction is infinitesimally small. Alternatively, we define an ESS in the case that the number of jurisdictions is finite according to Schaffer (1988).

Definition 1 (Finite population ESS) *A strategy τ^{ESS} is evolutionarily stable if for any other strategy τ,*

$$V\left(\tau^{ESS}, \tau + (n-2)\,\tau^{ESS}\right) > V\left(\tau, (n-1)\,\tau^{ESS}\right) \text{ for every } i \in J. \tag{18}$$

Let a jurisdiction choosing τ^{ESS} be called an *ESS strategist* while any $\tau \neq \tau^{ESS}$ be an *mutant strategist*. Definition 1 states that if $n-1$ ESS strategists can repel 1 mutant strategist, then τ^{ESS} is an evolutionarily stable strategy.

An ESS defined in Definition 1 can repel only one mutant strategy but may not be able to repel several mutant strategies once two or more identical mutant strategists appear simultaneously. More generally, in a state where m players choose $\tau \neq \tau^{ESS}$ and $n-m$ players τ^{ESS}, an ESS can repel m identical mutant strategist if $m \leq M$ but the mutant strategies can survive if $m > M$, where M is some integer between 0 and n. This implies that the greater the value of M is, the more simultaneous mutations can be repelled by the ESS. Therefore, the value of M means the degree of stability of the ESS in our tax competition game.

Definition 2 (Stability of ESS) *A strategy τ^{ESS} is M-stable if for any other strategy τ and for all $m \in [1, M]$ where $M \leq n-1$,*

$$V\left(\tau^{ESS}, m\tau + (n-m-1)\,\tau^{ESS}\right) > V\left(\tau, (m-1)\,\tau + (n-m)\,\tau^{ESS}\right), \tag{19}$$

for every $i \in J$. The ESS is globally stable if $M = n - 1$.

For the purpose of identifying an ESS, it is useful to consider the following relative utility:

$$\Pi\left(\tau, \tilde{\tau}; m\right) \equiv V\left(\tau, (m-1)\tau + (n-m)\tilde{\tau}\right) - V\left(\tilde{\tau}, m\tau + (n-m-1)\tilde{\tau}\right). \tag{20}$$

If $\tau = \tau^{ESS}$ is a strict global maximizer of $\Pi\left(\tau, \tau^{ESS}; m\right)$ for any $m \in [1, M]$, then τ^{ESS} is an ESS which is M-stable since $\Pi\left(\tau^{ESS}, \tau^{ESS}; m\right) = 0$ and thus $\Pi\left(\tau, \tau^{ESS}; m\right) < 0$ for all $\tau \neq \tau^{ESS}$. Suppose that τ^{ESS} is not a strict global maximizer. Then, there is some τ' such that $\Pi\left(\tau', \tau^{ESS}; m\right) \geq 0$. This means that $\tau = \tau^{ESS}$ is not an M-stable ESS. Hence, there exists a M-stable ESS if and only if $\Pi\left(\tau^{ESS}, \tau^{ESS}; m\right)$ is a strict global maximum for any $m \in [1, M]$. In order for τ^{ESS} to be a global maximizer, local maximizing conditions must be satisfied with $\tau = \tau^{ESS}$, i.e., $\frac{\partial}{\partial \tau}\Pi\left(\tau^{ESS}, \tau^{ESS}; m\right) = 0$ and $\frac{\partial^2}{\partial \tau^2}\Pi\left(\tau^{ESS}, \tau^{ESS}; m\right)$ < 0. In other words, if τ^{ESS} does not satisfy the local maximizing conditions, then it is not an ESS.

The first-order condition for the local maximum of $\Pi\left(\tau, \tau^{ESS}; m\right)$ is

$$\frac{\partial}{\partial \tau}\Pi\left(\tau^{ESS}, \tau^{ESS}; m\right) = \left(\bar{k} - \frac{\tau^{ESS}}{b}\right) u_g^i - \bar{k} u_x^i = 0, \tag{21}$$

or equivalently, $mrs^i = \frac{\bar{k}}{\bar{k} - \tau^{ESS}/b}$. Therefore, τ^{ESS} must be equal to τ^* satisfying (17) if it is a local maximizer. We have already seen that the symmetric Nash equilibrium strategy τ^{SNE} is greater than τ^* which is the symmetric Nash equilibrium strategy in the purely competitive case. This implies that τ^{SNE} cannot be a local maximizer of $\Pi\left(\tau, \tau^{SNE}; m\right)$ whatever m we have. Thus, we obtain the following proposition:

Proposition 2 *Let assumptions 1–3 hold and the production functions be quadratic. When there are finite number of jurisdictions, the unique symmetric Nash equilibrium strategy is not evolutionarily stable. If an ESS exists, then it is unique and satisfies (17).*

The uniqueness of the ESS is straightforward from the uniqueness of the symmetric Nash equilibrium in the purely competitive case since $\tau^{ESS} = \tau^*$. The unique ESS does not depend on the number of jurisdictions $n \geq 2$. That is, fiscal centralization through

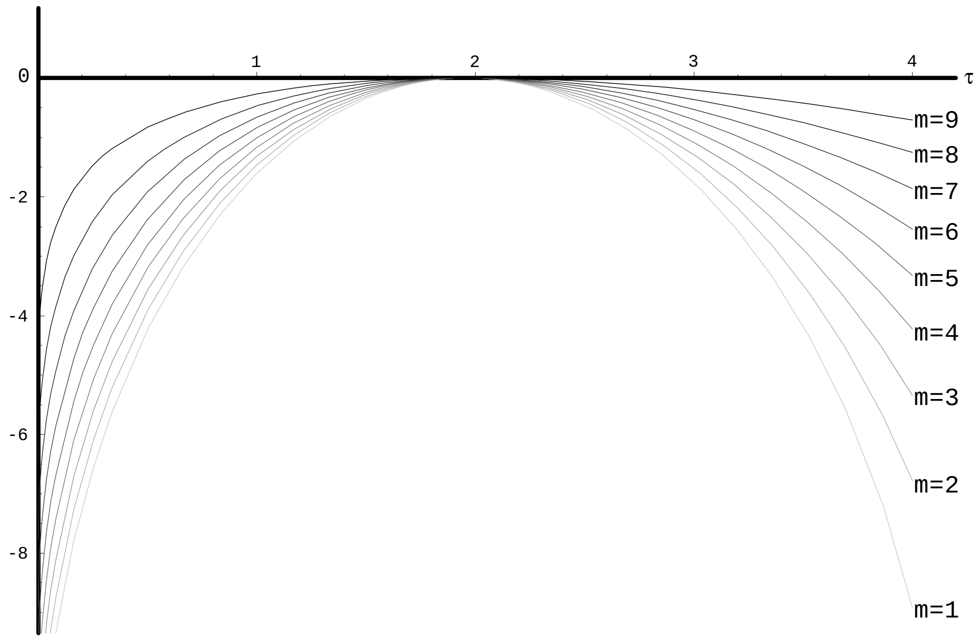

Figure 1: Relative utility curves and the globally ESS

a decrease in the number of jurisdictions could not reduce the efficiency losses of tax competition at all unless the fiscal system were completely centralized.

Since an ESS is required to be a global maximizer of $\Pi\left(\tau, \tau^{ESS}; m\right)$, it is very difficult to prove the existence of an ESS without the following assumption regarding preference in addition to assumptions 1–3:

Assumption 4. For every i and for all $x_i, g_i > 0$, $u^i_{gx} \equiv \frac{\partial^2}{\partial x_i \partial g_i} u\left(x_i, g_i\right) \geq 0$.

Although assumption 4 is stronger than assumptions 1–3, many specific utility functions such as the Cobb–Douglas utility and the CES utility satisfy assumption 4.

Proposition 3 *Let assumptions 1–4 hold and the production functions be quadratic. Then, there exists a unique global ESS that satisfies (17).*

Proof. See Appendix. ■

Figure 1 illustrates relative utility curves in numerical examples where $n = 10$, $a = 101$, $b = 2$ and $\bar{k} = 1$, and each individual has the Cobb–Douglas utility $u_i = x_i^{0.5} g_i^{0.5}$. In the purely competitive equilibrium of this example, we obtain $\tau^* \approx 1.96$. The relative utility curve for each $m \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ drawn in Figure 1 is globally maximized at $\tau^* \approx 1.96$. That is, the purely competitive equilibrium tax-rate $\tau^* \approx 1.96$ is equal to a unique global ESS.

5 Imitative learning dynamics

Evolutionary stability does not specify a dynamic selection process which forces to attain the state consisting of all the ESSs. Moreover, we cannot estimate the speed of convergence to a evolutionarily stable set as a long-run equilibrium. In this section, we explicitly model a situation where imitative behaviors among voters prevail in evolutionary dynamics. For technical reasons, we assume that the strategy space is a finite grid, i.e., $\Gamma = \{0, \Delta, 2\Delta, ..., z\Delta\}$ where $\Delta \in \mathbb{R}_{++}$ and $z \in \mathbb{N}$. Δ can be arbitrary small. The only restriction imposed on Γ is that $\tau^* \in \Gamma$. All that voters can observe are the capital tax rates $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)$, the private consumptions $\mathbf{x} = (x_1, \ldots, x_n)$, and the local public expenditures $\mathbf{g} = (g_1, \ldots, g_n)$. It is not plausible that each voter can observe utility levels of voters in other jurisdictions. Alternatively, the voters in jurisdiction i could know their own utility levels that would be attained if they consumed an observed bundle $\{x_j, g_j\}$.

The evolutionary dynamics proceed in discrete time, indexed by $t = 0, 1, 2, \ldots$. Denote the private consumption of the median voter i at t by $x_i(t)$, and the local public expenditure and the capital tax rate of jurisdiction i at t by $g_i(t)$ and $\tau_i(t)$, respectively. For brevity, we denote $V\left(\tau_i(t), \sum_{j \neq i} \tau_j(t)\right)$ by $v^i(t)$. At each t, there are two stages at which each government makes its decisions faithfully according to the median voter's demand. In the first stage, it is determined whether government i revises its tax rate $\tau_i(t-1)$ with a common and independent probability $\delta > 0$ or not. If government i decides to revise, the decision process proceeds to the second stage in which government i chooses from the set,

$$IM(t-1) = \left\{\tau_j(t-1) \in \Gamma \mid v^j(t-1) = \max\left\{v^l(t-1)\right\}_{l \in J}\right\}, \tag{22}$$

according to an independent probability distribution with full support. We can see that through the imitation rule defined in (22), the n-dimensional vector of tax rates at t, denoted by $\boldsymbol{\tau}(t)$, is determined by a n-dimensional vector of tax rates at $t-1$ $\boldsymbol{\tau}(t-1)$. That is, we have a discrete-time Markov process with finite state space Γ^n. Let a monomorphic state in which all governments choose the same capital tax rate τ be denoted by $mon(\tau)$ $= (\tau, \tau, \ldots, \tau)$. We obtain the following results:

Lemma 2 *Any monomorphic state $mon(\tau)$ is a singleton absorbing set of the imitative dynamics. And, any absorbing set is in $\{mon(\tau)\}_{\tau\in\Gamma}$.*

Proof. It is straightforward from the imitation rule that for any $\tau \in \Gamma$, the state $mon(\tau)$ leads to a absorbing set of the imitative dynamics. Suppose, contrary to our latter claim, that a absorbing set is in a set of nonmonomorphic states. Since each government chooses any $\tau \in IM\,(t-1)$ with positive probability, there is always positive probability that the process will transit from the set of nonmonomorphic states to the other set of states. This contradicts to our assumption that the set of nonmonomorphic states contains a absorbing set. ■

Therefore, there are $z+1$ singleton absorbing sets since the number of monomorphic states is equal to the number of the available tax rates in the finite strategy space Γ.

5.1 Stochastically stable states

We now introduce experimentation that will occur with some common independent probability $\epsilon > 0$. Once experimentation occurs, $\tau_i\,(t) \in \Gamma$ is chosen according to some given probability distribution with full support on Γ. Thus, for each $\epsilon > 0$, the perturbed Markov process has the irreducible transition matrix so that there is a unique invariant distribution ζ_ϵ which is independent of initial conditions and assigns positive probability to all states in Γ^n. The invariant distribution ζ_ϵ, however, depends on the occurrence of experimentation ϵ, and moreover will concentrate almost all of its probability on a few states as $\epsilon \to 0$ (see, Foster and Young [1990], Kandori, Mailath and Rob [1993], and Young [1993]). That is, we can investigate the relative robustness among the monomorphic states as the occurrence of experimentation vanishes. Hence, we focus on the limit invariant distribution of the perturbed Markov process as $\epsilon \to 0$, denoted by $\zeta^* \equiv \lim_{\epsilon\to 0} \zeta_\epsilon$. If a state is assigned some positive probability according to the limit invariant distribution ζ^*, then the state is called *stochastically stable*. A *stochastically stable set* consists of all states with positive probabilities. Stochastically stable states are, intuitively, the states that are most likely to be observed over the long run when the occurrence of experimentation is rare.

According to recent evolutionary literature, we use the techniques given by Freidlin and Wentzell (1984) to find a stochastically stable state. Any ordered pair of states is called "arrow", which is denoted by $(\boldsymbol{\tau}', \boldsymbol{\tau}'')$ for $\boldsymbol{\tau}', \boldsymbol{\tau}'' \in \Gamma^n$. For each $\boldsymbol{\tau} \in \Gamma^n$, a $\boldsymbol{\tau}$-tree is a collection of the arrows such that every $\boldsymbol{\tau}' \in \Gamma^n \setminus \{\boldsymbol{\tau}\}$ is the first element of the arrow, and for every $\boldsymbol{\tau}' \in \Gamma^n \setminus \{\boldsymbol{\tau}\}$ there is a path $\{(\boldsymbol{\tau}^0, \boldsymbol{\tau}^1), (\boldsymbol{\tau}^1, \boldsymbol{\tau}^2), \ldots, (\boldsymbol{\tau}^{s-1}, \boldsymbol{\tau}^s)\}$ where $\boldsymbol{\tau}^0 = \boldsymbol{\tau}'$ and $\boldsymbol{\tau}^s = \boldsymbol{\tau}$. The cost of the arrow $(\boldsymbol{\tau}', \boldsymbol{\tau}'')$ is the minimal number of experimentation needed for the transition from $\boldsymbol{\tau}'$ to $\boldsymbol{\tau}''$ to occur with positive probability. The cost of a $\boldsymbol{\tau}$-tree is the sum of the costs of all the arrows which belong to the $\boldsymbol{\tau}$-tree. The least cost among all $\boldsymbol{\tau}$-trees is the *stochastic potential* of $\boldsymbol{\tau}$. Young (1993) proves that the limit invariant distribution ζ^* assigns positive probability only to the states having minimum stochastic potential. That is, any state with a minimal-cost tree among all trees of all states is stochastically stable and a stochastically stable set consists of all the states with minimal-cost trees.

Proposition 4 *For capital tax competition among the finite number of jurisdictions under assumptions 1–4, there exists a unique stochastically stable state* $mon(\tau^*)$ *in the imitative dynamic.*

Proof. Since τ^* is the globally ESS, if $m = n - 1$, then using (20) we obtain that for every $\tau \in \Gamma \setminus \{\tau^*\}$,

$$V(\tau^*, (n-1)\tau) > V(\tau, (n-2)\tau + \tau^*). \tag{23}$$

The cost of $mon(\tau^*)$-tree is z since the inequality (23) implies that for every $\tau \in \Gamma \setminus \{\tau^*\}$ the total cost of the arrows along with a path from $mon(\tau)$ to $mon(\tau^*)$ is just one. Therefore, the stochastic potential of $mon(\tau^*)$ is equal to z. Clearly, z is the minimum stochastic potential. Hence, $mon(\tau^*)$ is contained in the stochastically stable set.

Next, we must show that the stochastically stable set is singleton, i.e., $mon(\tau^*)$ is a unique stochastically stable state. Lemma 2 states that no nonmonomorphic state can be included in the absorbing sets of the unperturbed dynamics, so that it cannot be stochastically stable since every stochastically stable state must be a absorbing state of the unperturbed dynamics.[5] Therefore, we have only to compare the cost of $mon(\tau^*)$-tree with the cost of the tree rooted in each of all the other monomorphic states. If m

[5] By continuity of ς_ϵ in ϵ, the limit invariant distribution of the perturbed dynamics ς^* is an invariant distribution of the unperturbed dynamics.

$= 1$, then we have from (20) that for every $\tau \in \Gamma \setminus \{\tau^*\}$,

$$V(\tau^*, \tau + (n-2)\tau^*) > V(\tau, (n-1)\tau^*). \tag{24}$$

This implies that given any $\tau' \in \Gamma \setminus \{\tau^*\}$, any transition from $mon(\tau^*)$ to $mon(\tau')$ does not occur with only one experimentation. That is, for every $\tau' \in \Gamma \setminus \{\tau^*\}$ the total cost of the arrows along with any path from $mon(\tau^*)$ to $mon(\tau')$ must be more than one, thus the cost of $mon(\tau')$-tree is more than z. Hence, for all $\tau' \in \Gamma \setminus \{\tau^*\}$, $mon(\tau')$ is not contained in the stochastically stable set. ■

Whenever the global ESS exists, a stochastically stable state is uniquely determined. In addition, the strategy in the unique SSS coincides with the global ESS. Since $\tau^* < \tau^{SNE}$, the unique SSS involves a greater efficiency loss than the Nash equilibrium. In other words, the imitative behaviors among boundedly rational voters tend to be harmful to allocative efficiency as compared with Nash behaviors by rational and well-informed voters. Moreover, fiscal centralization through a decrease in the number of jurisdictions could not reduce the efficiency losses of tax competition at all unless the fiscal system were completely centralized.

5.2 Speed of convergence

In the proof of Proposition 4, we have seen that the cost of the tree rooted in the unique stochastically stable state $mon(\tau^*)$ is one. This means that the basin of attraction of $mon(\tau^*)$, denoted by B^*, is the set consisting of all states where one or more jurisdictions choose τ^*.[6] The number of periods spent before arriving at $mon(\tau^*)$ starting from any initial state in B^* is hardly affected by the experimentation rate ϵ although it depends on the degree of inertia $1 - \delta$. Therefore, given any initial state in B^*, the time spent until reaching $mon(\tau^*)$ is reduced when δ is sufficiently close to one. On the other hand, in order to enter into B^* starting from any state outside B^*, at least one jurisdiction has to experiment and randomly choose τ^* from among $z + 1$ tax rates which can be chosen. Let us suppose that once experimentation occurs, a tax rate is randomly chosen from Γ according to the uniform distribution. Hence, the probability that each jurisdiction will choose τ^* at any state outside B^* is $\epsilon / (z+1)$ so that the expected number of jurisdictions

[6] The basin of attraction of $mon(\tau^*)$ is the set of initial states from which the unperturbed Markov process converges to $mon(\tau^*)$ with a probability of one.

Table 1: The expected number of periods spent outside B*

$z = 30$	ϵ				$z = 50$	ϵ			
n	.01	.05	.10	.20	n	.01	.05	.10	.20
3	1,034	207	103	52	3	1,700	340	170	85
5	620	124	62	31	5	1,020	204	102	51
10	310	62	31	16	10	510	102	51	26
20	155	31	16	8	20	255	51	26	13
30	104	21	11	6	30	170	34	17	9
40	78	16	8	4	40	128	26	13	7
50	62	13	7	4	50	102	21	11	6
100	31	7	4	2	100	51	11	6	3

choosing τ^* is $n\epsilon / (z+1)$. Thus, the expected number of periods spent outside B^* is given by

$$W(\epsilon, n, z) \equiv \frac{z+1}{n\epsilon}. \tag{25}$$

If ϵ is small but significant, then the expected waiting time $W(\epsilon, n, z)$ depends on n and z as well as ϵ. The effect of the number of jurisdictions on the expected waiting time is especially noteworthy in that we can compare with the comparative static results on the Nash equilibrium tax rate. The Nash equilibrium tax rate decreases with n and converges to τ^* as $n \to \infty$. In the imitative learning dynamics, $W(\epsilon, n, z)$ becomes shorter as n increases. That is to say, the higher the degree of fiscal decentralization, the faster convergence to $mon(\tau^*)$ starting at any initial state outside B^*. This result is derived from 'comparative dynamics' where the effect of n on the dynamic path is investigated.

In a fiscal competition model, it should be presumed that the time interval until a current policy can be revised is longer than in other dynamic models, e.g., oligopolistic competition where firms revise their outputs or prices in every period. When the number of periods needed to reach $mon(\tau^*)$ is very large, therefore, the system may not converge to $mon(\tau^*)$ within a reasonable time horizon. If it is necessary to wait for hundreds or thousands of years before convergence to $mon(\tau^*)$, we could not derive any meaningful policy implication from $mon(\tau^*)$ as a long-run equilibrium. Table 1 shows the results of numerical simulation for the expected waiting time before entering in B^*. When $z = 30$ and $n = 40$, 50 or 100, it can be expected to converge to $mon(\tau^*)$ within 100 periods for any $\epsilon \in \{.01, .05, .10, .20\}$ even though the initial state is outside B^*. In the case that $z = 30$ and $n = 5$, on the other hand, the number of periods needed for convergence would

Table 2: The ratio of time spent outside B* to time spent in B*

$z = 30$	ϵ			
n	.01	.05	.10	.20
2	.1452	.7258	1.4516	2.9032
3	.0009	.0234	.0937	.3746
5	–	–	.0005	.0084
10	–	–	–	–
$z = 50$	ϵ			
n	.01	.05	.10	.20
2	.2451	1.2255	2.4511	4.9021
3	.0016	.0401	.1602	.6408
5	–	–	.0009	.0148
10	–	–	–	–

Note: A dash indicates less than .00006.

be more than 100 for each $\epsilon = .01$ and .05. If we cannot regard 100 periods and more as a reasonable time horizon, then for $z = 30$ and $n = 5$ $mon\,(\tau^*)$ would not be a good prediction when $\epsilon = .01$ and .05 but when $\epsilon = .10$ and .20. In general, from (25), the smaller n and/or the larger z, the higher ϵ is needed to hasten convergence to $mon\,(\tau^*)$.

A unique stochastically stable state is the only state that is assigned probability one as $\epsilon \to 0$. Intuitively, when the experimentation rate is very close to zero, the system spends almost all of its time in the long-run equilibrium where all jurisdictions choose τ^*. With a sizable experimentation rate (say, $\epsilon = .10$ or .20), how long does the system spend in B^* relative to the number of periods spent outside B^*? The probability that each jurisdiction will experiment and choose any τ except for τ^* in B^* is $\epsilon z/(z+1)$ and the system needs n simultaneous experimentations to leave B^*. Therefore, the probability that the system will leave B^* at any state in B^* is given by $[\epsilon z/(z+1)]^n$, and we have $[\epsilon z/(z+1)]^{-n}$ which is the expected number of periods spent in B^*. Thereby, using (25), the ratio of time spent outside B^* to time spent in B^* is

$$R\,(\epsilon, n, z) \equiv \frac{z^n}{n\,(z+1)^{n-1}}\epsilon^{n-1}. \tag{26}$$

Since R converges to zero as $\epsilon \to 0$, for a sufficiently small ϵ the system spends most of its time in B^* and therefore at the stochastically stable state $mon\,(\tau^*)$.[7] The values of R

[7]This property is consistent with the radius-coradius theorem in Ellison (2000). The radius-coradius theorem states that if for some set Ω that is a union of limit sets the radius of Ω is greater than the

in (26) for each $\epsilon \in \{.01, .05, .10, .20\}$ are summarized in Table 2. If there are 5 or more jurisdictions, the system would spend almost of its time in B^* for every selected value of ϵ.[8] Whereas, in the case that $n = 3$, for each $\epsilon = .05$, .10 and .20 we could not regard that the time spent outside B^* is negligible although the time spent in B^* is relatively long. Furthermore, when $n = 2$, there are some cases in which the time spent outside B^* exceeds the time in B^*. For $\epsilon = .001$, $n = 2$ and $z = 30$, we have $R = .0145$ while it will be necessary to wait for $15,500$ periods before entering in B^*. To sum up, for a sufficiently large n the system spends most of its time in the stochastically stable state $mon\,(\tau^*)$ even with a fairly large value of ϵ by which the system will converge to $mon\,(\tau^*)$ within 100 periods.

6 Concluding remarks

Both the unique globally ESS and the unique SSS have the same result as the purely competitive model of capital tax competition. These results are consistent with the results of Schaffer(1989) and Vega-Redondo(1997). A static interpretation of these long-run equilibria is that regional decision makers choose their tax rate taking the economy-wide net return to capital as exogenously given.[9] It should be noted that the firms behave as if their choices cannot affect the other jurisdictions' tax revenues at all in tax competition among finite and nonatomistic jurisdictions. Hoyt(1991) argues that the optimal number of jurisdictions can be determined if a trade-off between gains from the sorting of residents by voting with their feet and the costs of capital tax competition is considered. Our result implies that fiscal centralization is preferable to any decentralized system if the costs of capital tax competition are greater than the gains from Tiebout sorting, and otherwise the number of jurisdictions should be increased as much as possible.

Because selection dynamics behind evolutionary stability are not clear, we can give a dynamic interpretation of the ESS in the context of our tax competition model. Suppose that tax policies are chosen under a representative democracy. Representatives may be rational and have enough information but they are concerned with winning the next

coradius of Ω, then the stochastically stable set is contained in Ω. In our model, the radius of B^* is n while the coradius of B^* is 1.

[8] For any $\epsilon<1$ and z, we can easily verify that R decreases with n.

[9] The results in Schaffer(1989) and Vega-Redondo(1997) are both a Cournot–Nash equilibrium as firms choose their outputs taking a price as exogenously given.

election. Therefore, their political decisions could be influenced by pressures from voters who prefer the tax policy most successful among all jurisdictions. Consequently, the objective of representatives that survives economic natural selection is that of maximizing the relative utility given by (20) rather than maximizing the absolute utility of the representative voter in their own jurisdiction. Thus, the idea of evolutionary stability provides an insight into the objective of regional politicians.

Appendix

Proof of Lemma 1. The first-order derivative of $V\left(\tau_i, \sum_{j\neq i}\tau_j\right)$ with respect to τ_i is

$$\begin{aligned} A_i &\equiv \left(k_i - \frac{n-1}{n}\frac{\tau_i}{b}\right)u_g^i - \frac{1}{n}\left[(n-1)\,k_i + \bar{k}\right]u_x^i \\ &= \left(2u_g^i - \frac{n-1}{n}u_x^i\right)k_i - \left(k_i + \frac{n-1}{n}\frac{\tau_i}{b}\right)u_g^i - \frac{\bar{k}}{n}u_x^i, \end{aligned} \tag{A1}$$

when $k_i = \bar{k} + (\bar{\tau} - \tau_i)/b$. Differentiating A_i with respect to τ_i yields

$$\frac{\partial A_i}{\partial \tau_i} = -\left(2u_g^i - \frac{n-1}{n}u_x^i\right)\frac{1}{nb} + \left(k_i - \frac{n-1}{n}\frac{\tau_i}{b}\right)^2\Phi, \tag{A2}$$

where

$$\Phi \equiv u_{gg}^i - 2u_{gx}^i\mu_i + u_{xx}^i\mu_i^2, \tag{A3}$$

together with

$$\mu_i \equiv \frac{\frac{1}{n}\left[(n-1)\,k_i + \bar{k}\right]}{k_i - \frac{n-1}{n}\frac{\tau_i}{b}}. \tag{A4}$$

We will show that (A2) is negative when $A_i \geq 0$, that is, $V\left(\tau_i, \sum_{j\neq i}\tau_j\right)$ is strictly concave (i.e., $\partial A/\partial\tau_i < 0$) unless it is decreasing. This implies that $V\left(\tau_i, \sum_{j\neq i}\tau_j\right)$ has a single peak with respect to τ_i. The first term of (A2) is negative since $2u_g^i - \frac{n-1}{n}u_x^i$ must be positive when $A_i \geq 0$ from (A1). The numerator in (A4) is positive since $k_i > 0$. In addition, since $k_i - \frac{n-1}{n}\frac{\tau_i}{b} > 0$ from (A1) whenever $A_i \geq 0$, we have $\mu_i > 0$. Thus, if $u_{gx}^i \geq 0$, then $\Phi \leq 0$ since $u_{gg}^i \leq 0$ and $u_{xx}^i \leq 0$ (Assumption 1), which leads to $\partial A/\partial\tau_i < 0$. Consider the case that $u_{gx}^i < 0$. From Assumption 2, the marginal rate of substitution u_g^i/u_x^i is nondecreasing in x_i and nonincreasing in g_i. Therefore, we have $u_{gx}^i u_x^i - u_{xx}^i u_g^i$

≥ 0 and $u^i_{gg}u^i_x - u^i_{gx}u^i_g \leq 0$, or equivalently $\frac{u^i_{gg}}{u^i_{gx}} \geq \frac{u^i_g}{u^i_x} \geq \frac{u^i_{gx}}{u^i_{xx}}$. We can easily verify that $\Phi \leq 0$ for any μ_i when $\frac{u^i_{gg}}{u^i_{gx}} \geq \frac{u^i_{gx}}{u^i_{xx}}$. Hence, we also have $\partial A/\partial \tau_i < 0$ when $u^i_{gx} < 0$. ■

Proof of Proposition 3. Let the amounts of capital employed in each jurisdiction choosing τ and τ^*, respectively, be denoted by k^+ and k^-, respectively. Using (4), we have

$$k^+ = \bar{k} + \frac{(n-m)(\tau^* - \tau)}{nb}, \; k^- = \bar{k} + \frac{m(\tau - \tau^*)}{nb}. \tag{A5}$$

Correspondingly, we denote the private consumptions x^+ and x^-, and the local public services g^+ and g^-.

$$x^+ = \frac{b}{2}\left[\bar{k} + \frac{(n-m)(\tau^* - \tau)}{nb}\right]^2 - \frac{m\tau + (n-m)\tau^*}{n}\bar{k} + \left(a - b\bar{k}\right)\bar{k}, \tag{A6a}$$

$$x^- = \frac{b}{2}\left[\bar{k} + \frac{m(\tau - \tau^*)}{nb}\right]^2 - \frac{m\tau + (n-m)\tau^*}{n}\bar{k} + \left(a - b\bar{k}\right)\bar{k}, \tag{A6b}$$

$$g^+ = \tau\left[\bar{k} + \frac{(n-m)(\tau^* - \tau)}{nb}\right], \tag{A6c}$$

$$g^- = \tau^*\left[\bar{k} + \frac{m(\tau - \tau^*)}{nb}\right]. \tag{A6d}$$

Thus, we obtain

$$x^+ - x^- = \left(k^+ + k^-\right)\frac{\tau^* - \tau}{2} \lesseqgtr 0 \text{ if } \tau \gtreqless \tau^*, \tag{A7a}$$

$$g^+ - g^- = (\tau - \tau^*)\left[\bar{k} - \frac{m\tau^* + (n-m)\tau}{nb}\right] \gtrless 0 \quad \begin{array}{c}\text{if } \tau^* < \tau < \hat{\tau}, \\ \text{if } \tau < \tau^* \text{ or } \tau > \hat{\tau},\end{array} \tag{A7b}$$

where $\hat{\tau} \equiv \left(nb\bar{k} - m\tau^*\right)/(n-m) > \tau^*$.

Differentiating (A6a)-(A6d) with respect to τ yields

$$\frac{\partial x^+}{\partial \tau} = -\frac{n-m}{n}\left[\bar{k} + \frac{(n-m)(\tau^* - \tau)}{nb}\right] - \frac{m\bar{k}}{n} < 0, \tag{A8a}$$

$$\frac{\partial x^-}{\partial \tau} = \frac{m^2}{n^2 b}(\tau - \tau^*) \gtreqless 0 \text{ if } \tau \gtreqless \tau^*, \tag{A8b}$$

$$\frac{\partial g^+}{\partial \tau} = \bar{k} + \frac{(n-m)(\tau^* - 2\tau)}{nb} \lesseqgtr 0 \text{ if } \tau \gtreqless \frac{nb\bar{k} + (n-m)\tau^*}{2(n-m)} \equiv \check{\tau}, \tag{A8c}$$

$$\frac{\partial g^-}{\partial \tau} = \frac{m\tau^*}{nb} > 0. \tag{A8d}$$

Letting $mrs\left(x^{+},g^{+}\right) \equiv u_g\left(x^{+},g^{+}\right)/u_x\left(x^{+},g^{+}\right)$, differentiating $mrs\left(x^{+},g^{+}\right)$ with respect to τ yields

$$\frac{\partial}{\partial \tau} mrs\left(x^{+},g^{+}\right) = \frac{\partial mrs}{\partial g^{+}}\frac{\partial g^{+}}{\partial \tau} + \frac{\partial mrs}{\partial x^{+}}\frac{\partial x^{+}}{\partial \tau}. \tag{A9}$$

From (A8a) and (A8c) together with Assumption 2, $mrs\left(x^{+},g^{+}\right)$ decreases with τ as long as τ is lower than or equal to $\check{\tau}$. Moreover, differentiating $\Pi\left(\tau,\tau^{*};m\right)$ with respect to τ, we obtain

$$\frac{\partial}{\partial \tau}\Pi\left(\tau,\tau^{*};m\right) = u_x\left(x^{+},g^{+}\right)\pi\left(\tau\right) + \frac{m}{n}\psi, \tag{A10}$$

where

$$\begin{aligned} \pi\left(\tau\right) &\equiv \left(k^{+} - \frac{\tau}{b}\right) mrs\left(x^{+},g^{+}\right) - k^{+} \\ &= k^{+}\left[mrs\left(x^{+},g^{+}\right) - 1\right] - \frac{\tau}{b} mrs\left(x^{+},g^{+}\right), \qquad \text{(A11a)} \\ \psi &\equiv \frac{\tau}{b}u_g\left(x^{+},g^{+}\right) - \frac{\tau^{*}}{b}u_g\left(x^{-},g^{-}\right) \\ &\quad + \left(k^{+}-\bar{k}\right)u_x\left(x^{+},g^{+}\right) - \left(k^{-}-\bar{k}\right)u_x\left(x^{-},g^{-}\right). \qquad \text{(A11b)} \end{aligned}$$

We consider three cases with the range in which τ can be taken. Because it is ambiguous whether $\check{\tau}$ is greater or lower than $\hat{\tau}$, we use $\sigma \equiv \min\{\hat{\tau},\check{\tau}\}$.

Case 1 ($\tau < \tau^{*}$) *Differentiating $\pi\left(\tau\right)$ with respect to τ yields*

$$\frac{\partial \pi}{\partial \tau} = \left(k^{+} - \frac{\tau}{b}\right)\frac{\partial mrs}{\partial \tau} - \frac{n-m}{nb}\left[mrs\left(x^{+},g^{+}\right) - 1\right] - \frac{mrs}{b}. \tag{A12}$$

Since $mrs\left(x^{+},g^{+}\right)$ is decreasing in $\tau \leq \check{\tau}$, $mrs\left(x^{+},g^{+}\right) \geq mrs\left(x^{*},g^{*}\right) > 1$ when $\tau \leq \tau^{*}$. Moreover, $k^{+} - \tau/b \geq \bar{k} - \tau^{*}/b > 0$ since $k^{+} \geq \bar{k}$ when $\tau \leq \tau^{*}$. Thus, we have $\partial\pi/\partial\tau < 0$. Note that $\pi\left(\tau^{*}\right) = \left(\bar{k} - \tau^{*}/b\right) mrs\left(x^{*},g^{*}\right) - \bar{k} = 0$. Hence, we obtain $\pi\left(\tau\right) > 0$ for any $\tau < \tau^{*}$ which yields $\frac{\partial}{\partial\tau}\Pi > 0$ for any $\tau < \tau^{*}$ if $\psi \geq 0$. Suppose that $\psi < 0$. We can rewrite (A10) as

$$\begin{aligned} \frac{\partial}{\partial \tau}\Pi\left(\tau,\tau^{*};m\right) = &\left(\bar{k} - \frac{\tau^{*}}{b}\right)u_x\left(x^{+},g^{+}\right)\left[mrs\left(x^{+},g^{+}\right) - \frac{\bar{k}}{\bar{k}-\tau^{*}/b}\right] \\ &- \frac{n-m}{n}\psi + \left(k^{+}-\bar{k}\right)u_g\left(x^{+},g^{+}\right) + \left(\bar{k}-k^{-}\right)u_x\left(x^{-},g^{-}\right) \\ &+ \frac{\tau^{*}}{b}\left[u_g\left(x^{+},g^{+}\right) - u_g\left(x^{-},g^{-}\right)\right]. \end{aligned} \tag{A13}$$

Since $mrs\left(x^{+},g^{+}\right) > mrs\left(x^{*},g^{*}\right) = \frac{\bar{k}}{\bar{k}-\tau^{*}/b} > 0$*, the first term of (A13) is positive. As we assume that* $\psi < 0$*, the second term is also positive. The third and the forth terms are both positive since* $k^{+} > \bar{k}$ *and* $\bar{k} > k^{-}$*, respectively. Since* $x^{+} > x^{-}$ *and* $g^{+} < g^{-}$ *from (A7a) and (A7b), using the assumptions that* $u_{gg} \leq 0$ *and* $u_{gx} \geq 0$*, we have*

$$\begin{aligned} 0 &= u_g\left(x^{+},g^{+}\right) - u_g\left(x^{+},g^{+}\right) \\ &\leq u_g\left(x^{+},g^{+}\right) - u_g\left(x^{-},g^{+}\right) \\ &\leq u_g\left(x^{+},g^{+}\right) - u_g\left(x^{-},g^{-}\right). \end{aligned} \quad \text{(A14)}$$

Thus, the fifth term of (A13) is nonnegative. Hence, $\Pi\left(\tau,\tau^{*};m\right)$ *is strictly increasing in* $\tau < \tau^{*}$ *regardless of whether* ψ *is positive or negative.*

Case 2 $(\sigma \geq \tau > \tau^{*})$ *From (A11a),* $\pi\left(\tau\right) < 0$ *for any* $\tau \in \left(\tau^{*},\check{\tau}\right]$ *if* $k^{+} - \tau/b \leq 0$ *or* $mrs\left(x^{+},g^{+}\right) - 1 \leq 0$*. Consider any* $\tau \in \left(\tau^{*},\check{\tau}\right]$ *such that* $k^{+} - \tau/b > 0$ *and* $mrs\left(x^{+},g^{+}\right) - 1 > 0$*. Then, we have* $\pi\left(\tau\right) < 0$ *since* $\pi\left(\tau^{*}\right) = 0$ *and* $\partial\pi/\partial\tau < 0$ *for any* $\tau \in \left[\tau^{*},\check{\tau}\right]$ *from (A12). Thus,* $\Pi\left(\tau,\tau^{*};m\right)$ *has a negative slope when* $\psi \leq 0$*. Suppose that* ψ *is positive. Then, we can verify that the signs from the first to the forth term in (A13) are all negative and the fifth term in (A13) is nonpositive from the assumptions that* $u_{gg} \leq 0$ *and* $u_{gx} \geq 0$*. Hence, whenever* $\check{\tau} \geq \tau > \tau^{*}$, $\Pi\left(\tau,\tau^{*};m\right)$ *is strictly decreasing in* τ*.*

Case 3 $(\tau > \sigma > \tau^{*})$ *Suppose that* $\sigma = \check{\tau}$*. Then, we have* $\frac{\partial g^{+}}{\partial\tau} < 0$, $\frac{\partial x^{+}}{\partial\tau} < 0$, $\frac{\partial g^{-}}{\partial\tau} > 0$ *and* $\frac{\partial x^{-}}{\partial\tau} > 0$ *for* $\tau \geq \check{\tau}$*. Since* $\Pi\left(\check{\tau},\tau^{*};m\right) < 0$ *or* $u\left(x^{+},g^{+}\right) < u\left(x^{-},g^{-}\right)$ *when* $\tau = \check{\tau}$, $\Pi\left(\tau,\tau^{*};m\right) < 0$ *for any* $\tau \geq \check{\tau}$*. If* $\sigma = \hat{\tau}$*, then we have* $x^{+} < x^{-}$ *and* $g^{+} < g^{-}$ *which leads to* $u\left(x^{+},g^{+}\right) < u\left(x^{-},g^{-}\right)$ *or equivalently* $\Pi\left(\tau,\tau^{*};m\right) < 0$*.*

The results derived in Cases 1–3 imply that $\Pi\left(\tau,\tau^{*};m\right)$ is globally maximized at $\tau = \tau^{*}$. ■

References

Alós-Ferrer, C., Ania, A.B., Schenk-Hoppé, K.R., 2000. An evolutionary model of Bertrand oligopoly. Games and Economic Behavior 33, 1–19.

Bergin, J., Bernhardt, D., 2004. Comparative learning dynamics. International Economic Review 45(2), 431–465.

Besley, T., Case, A., 1995. Incumbent behavior: vote-seeking, tax-setting, and yardstick competition. American Economic Review 85(1), 25–45.

Bucovetsky, S., 1991. Asymmetric tax competition. Journal of Urban Economics 30, 167–181.

Ellison, G., 2000. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. Review of Economic Studies 67, 17–45.

Foster, D., Young, P., 1990. Stochastic evolutionary game dynamics. Theoretical Population Biology 38, 219–232.

Freidlin, M.I., Wentzell, A.D., 1984. Random perturbations of dynamic systems. Springer-Verlag, New York.

Hoyt, W.H., 1991. Property taxation, Nash equilibrium, and market power. Journal of Urban Economics 30, 123–131.

Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61(1), 29–56.

Shaffer, M.E., 1988. Evolutionarily stable strategies for a finite population and a variable contest size. Journal of Theoretical Biology 132, 469–478.

Shaffer, M.E., 1989. Are profit-maximizers the best survivors?: A Darwinian model of economic natural selection. Journal of Economic Behavior and Organization 12, 29–45.

Tiebout, C.M., 1956. A pure theory of local expenditures. Journal of Political Economy 64, 416–424.

Vega-Redondo, F., 1997. The evolution of Walrasian behavior. Econometrica 65(2), 375–384.

Wildasin, D.E., 1988. Nash equilibria in models of fiscal competition. Journal of Public Economics 35, 229–240.

Wilson, J.D., 1986. A theory of interregional tax competition. Journal of Urban Economics 19, 296–315.

Wrede, M., 2001. Yardstick competition to tame the Leviathan. European Journal of Political Economy 17, 705–721.

Young, P., 1993. The evolution of conventions. Econometrica 61(1), 57–84.

Zodrow, G.R., Mieszkowski, P., 1986. Pigou, Tiebout, property taxation, and the under-provision of local public goods. Journal of Urban Economics 19, 356–370.