Jean Hindriks and Yukihiro Nishimura

Taxing multinationals: The scope for enforcement cooperation

CORE

CENTER FOR
OPERATIONS RESEARCH
AND ECONOMETRICS

Taxing Multinationals: The Scope for Enforcement Cooperation[*]

Jean HINDRIKS[†] and Yukihiro NISHIMURA[‡]

September 25, 2018

Abstract

We present a tax-competition model with two policy instruments: the corporate tax rate and the tightness of tax enforcement (i.e., controls on profit shifting by multinational enterprises). Tougher enforcement increases the cost of profit shifting, and thus mitigates tax competition. We compare the equilibria of the noncooperative and cooperative enforcement choices. After showing that enforcement cooperation may not benefit the low-tax country, we indicate two drivers that promote enforcement cooperation. The first driver of cooperation is complementarity (imperfect substitutability) of countries' enforcement efforts, taking into account that dispersed enforcement efforts among the involved countries are less effective. We show that cooperation is more likely with greater enforcement complementarity. The second driver of cooperation is tax leadership, which reduces the extent of disagreement on tax enforcement.

Keywords: Profit shifting; Tax competition; Tax enforcement; Weakest-link; Tax leadership

JEL Classification: C72, F23, F68, H25, H87.

---

[*]**Declarations of interest: none.**

[†]CORE, Université catholique de Louvain. E-mail: `jean.hindriks@uclouvain.be`

[‡]Corresponding author. Graduate School of Economics, Osaka University, 1-7 Machikaneyama-cho, Toyonaka-shi, Osaka, 560-0043, Japan. Tel & fax: +81 (6)-6850-5230. E-mail: `ynishimu@econ.osaka-u.ac.jp`

1 Introduction

Globalization has prompted the emergence of multinational enterprises (MNEs) with divisions in different countries. Under the widely used system of source-based taxation (also known as a territorial system),[1] these MNEs can reduce their overall tax liabilities by shifting profit and income between branches, which causes the well-known problem of base erosion and profit shifting (BEPS, hereafter referred to as "profit shifting"). There is convincing empirical evidence of profit shifting.[2] To address this issue, the OECD's (2015) BEPS report proposed global action plans containing a series of specific enforcement efforts to be undertaken by governments. The European Union (EU) also established the Fiscalis 2020 program (European Commission, 2018) to ensure exchange of information and to support administrative cooperation. However, a real challenge to international cooperation is the absence of a global institution with enforcement powers. In reality, the enforcement of legislation differs significantly across countries: some countries only loosely acknowledge the "arm's length principle", whereas others ask firms to submit detailed transfer pricing reports for strict tax compliance purposes.

In this paper, we analyze the determinants of voluntary cooperation regarding enforcement efforts and information sharing arrangements. To address these issues in a tax competition framework, we develop a simple two-country model with different market sizes, following Kanbur and Keen (1993), Hindriks et al. (2014), and Keen and Konrad (2013). MNEs shift profits from the division in the high-tax country to that in the low-tax country, subject to a concealment cost. Countries choose their enforcement effort levels, which involve activities such as strict monitoring and inspection, more efficient information sharing, reinforcement of tax officials' skills and competence, and efforts to negotiate and reach agreements with the other country's tax authority. The enforcement efforts incur administration costs in each country. However, their benefit, which is the reduction of aggressive tax planning to mitigate the tax competition, is diffused to the other country. Therefore, choosing enforcement is analogous to making a voluntary contribution

---

[1]Most OECD countries have adopted territorial systems because it is relatively easy for taxpayers to avoid resident-based taxes on capital income. Profit shifting is also observed in the United States; see, for example, Swenson (2001), Clausing (2003), and Grubert and Altshuler (2013).

[2]For example, Mintz and Smart (2004) found that Canadian firms that operate in multiple jurisdictions have a high elasticity of taxable income with respect to corporate tax rates. Bartelsman and Beetsma (2003) used data from 22 OECD countries and found that more than 65 percent of the additional revenue resulting from a unilateral tax increase is lost because of profit shifting. See also Swenson (2001), Clausing (2003), Huizinga and Laeven (2008), and Næss-Schmidt et al. (2012).

to a public good. Asymmetric tax competition models with multiple policy instruments (tax rates and enforcement levels in this paper) are usually complex. However, by reducing the model to a few key components, we are able to explicitly solve it. We then compare the equilibria for the noncooperative and cooperative enforcement choices. In the latter scenario, countries choose enforcement levels to maximize their joint welfare, but they still determine tax rates noncooperatively.[3] With sufficient disparity in the countries' sizes, the low-tax country is not willing to cooperate on enforcement. This resistance to cooperation is of central concern in the application of the OECD's (2015) coordination proposals to the G20 and the EU because the adoption of such proposals often requires unanimity.

Given this context, we identify two different drivers of enforcement cooperation. The first driver is complementarity in the enforcement choices. Enforcement efforts increase the cost of profit shifting. However, in reality, if the countries undertake dispersed (unilateral) enforcement efforts, these are less effective. For instance, if the host country does not provide tax-relevant information, the home country's tax authorities cannot address tax fraud effectively. Indeed, in regard to the recommendations of the OECD BEPS project, including treatment of controlled foreign company rules (Action 3), interest deductions and other financial payments (Action 4), and mandatory disclosure of cross-border tax planning (Action 12), countries are free to choose whether to introduce them, based on each country's domestic tax regime.[4] Further, during a mutual agreement procedure on MNEs' taxable incomes, the low-tax country may favor the MNEs' transfer pricing methods (supported by their financial and legal experts) that result in the accrual of higher taxable incomes to the low-tax country. Given that double taxation is not allowed under the tax treaty, the low-tax country can in this way exercise a veto power on tighter enforcement, and the high-tax country has to be constrained by the minimum level of enforcement standards. This scenario corresponds to the "weakest–link" case analyzed by Hirshleifer (1983, p. 373). We formalize the collective-action problems of enforcement by using Hirshleifer's (1983) social composition function. Depending on types of enforcement contributions, different aggregation technology will prevail, from perfect substitutability to the weakest–link. We show that stronger complementarity lowers the noncooperative enforcement level and encourages profit shifting (Lemma 4). Such an enforcement structure negatively affects the tax competition game. As a result, enforcement cooperation is more likely

---

[3]The host country retains the primary taxing rights over source-based incomes. In the asymmetric tax competition models, countries have different preferred tax rates.

[4]Such weak forms of recommendation are called "Best Practices". Action 14 (effectiveness of dispute resolution) is appended with eleven Best Practices for implementation.

with greater enforcement complementarity (Proposition 1).

The second driver of cooperation is tax leadership, which has attracted research interest since Kempf and Rota-Graziosi (2010).[5] Following Hindriks and Nishimura (2015, 2017), we consider the case in which the large country leads.[6] Compared with simultaneous tax choice, tax leadership induces higher equilibrium taxes and reduces the extent of disagreement on enforcement (Lemma 6). As a result, enforcement cooperation will be more likely than under the Nash tax competition (Proposition 2).

The main focus of the previous studies was to understand how transfer prices are affected by international tax differences and tax systems.[7] In addition, some evidence showed that transfer pricing regulations significantly mitigate profit shifting (for example, Bartelsman and Beetsma (2003) and Lohse and Riedel (2013)). As in the current paper, several studies modeled tax competition augmented with enforcement choices. In the context of strategic information exchange, Bacchetta and Espinosa (1995) showed that countries do not choose the residence-based principle regarding foreigners' investment incomes. Peralta et al. (2006) showed that a country may adopt a lenient enforcement policy in equilibrium to tax-discriminate between the domestic firms and multinational firms, the latter of which can shift profit outside. Bucovetsky and Haufler (2008) showed that tougher enforcement reduces the possibility of tax discrimination between internationally mobile and immobile firms, and may intensify tax competition.[8] As such, the focus of their analysis is different from ours. In particular, none of these papers discussed the determinants of voluntary cooperation on enforcement efforts

---

[5]In the context of double-taxation conventions on capital income taxes, Gordon (1992) showed that capital income will be taxed in equilibrium if a dominant capital exporter acts as a Stackelberg leader. Altshuler and Goodspeed (2015) demonstrated that European countries set their corporate tax rates in line with the United States 1986 Tax Reform Act. Chatelais and Peyrat (2008) provided evidence of sequential tax decisions among European countries. Stöwhase (2013) adopted a Stackelberg setting simply to avoid the difficulties resulting from possible discontinuities in the payoff functions with a simultaneous tax setting. As such, in his framework, the Nash and the Stackelberg tax competitions are not comparable. Further, in contrast to our paper, the public-good nature of tax enforcement is absent in Stöwhase (2013).

[6]As we discuss later, the large leadership equilibrium Pareto dominates the small leadership equilibrium under sufficient asymmetry, as in Hindriks and Nishimura (2015).

[7]See, for example, Swenson (2001), Kind et al. (2005), Devereux et al. (2008), Nielsen et al. (2008), Huizinga and Laeven (2008), Amerighi and Peralta (2010), and Klassen and Laplante (2012).

[8]Konrad (2008) endogenized the size of the groups of mobile and immobile firms based on the loyalty (home attachment) of the citizens. A country's investment in citizen loyalty increases the tax revenue from loyal citizens, but it becomes a disadvantage in the tax competition game.

emerging. Cremer and Gahvari (2000) added enforcement decisions to the tax competition model and showed that tax harmonization could be damaging because it would induce countries to lower tax enforcement.[9] However, in contrast with Cremer and Gahvari (2000), in our model, the incomes shifted by an MNE from one country's branch to other countries' branches become the taxable incomes of other countries. As such, one country's enforcement efforts directly affect the tax base of other countries. Using spatial econometric approach, Durán-Cabré et al. (2015) confirmed positive horizontal interactions (strategic complementarities) between regional administrations in regard to audit policies among Spanish regional governments.

The rest of the paper is organized as follows. Section 2 describes the model and the timing of the game. Section 3 derives the equilibrium under noncooperative and cooperative enforcement regimes, and examines the benefit of cooperation in relation to enforcement complementarity. Section 4 examines tax leadership. Section 5 concludes. The proofs of propositions and lemmas are provided in the Appendices.

2 Framework

2.1 The model

The model used follows Hindriks et al. (2014) and Keen and Konrad (2013). There are two countries, denoted by 1 and 2. Each country has a linear (inverse) demand for a homogeneous good $p_i(q_i) = \alpha_i - \beta q_i$ $(i = 1, 2)$. Two multinational enterprises (MNEs), a and b, have branches in each country and compete *à la* Cournot in each domestic market. For each firm, production incurs the country-specific unit cost $c_i \geq 0$ $(i = 1, 2)$. We assume that:

$$\alpha_1 - c_1 \equiv \gamma_1 \geq \alpha_2 - c_2 \equiv \gamma_2 > 0.$$

That is, country 1 is the *large* country because (i) α_i, which represents population size or income per capita, is higher, or (ii) the supply cost c_i is lower.

From the production decisions (q_i^a, q_i^b), in country $i = 1, 2$, firm $k = a, b$ generates $\pi_i^k = \{p_i(q_i^a + q_i^b) - c_i\}q_i^k$ in country i. Then, at some cost, it may shift profits between branches to minimize the firm's total tax liability. In other words, it decides how much profit to report, $\tilde{\pi}_i^k$ in country i, where total reported profit must equal total realized profit $(\tilde{\pi}_1^k + \tilde{\pi}_2^k = \pi_1^k + \pi_2^k)$. Given country i's source-based tax rate t_i on the reported profit, firm k's profit becomes $(1-t_1)\tilde{\pi}_1^k + (1-t_2)\tilde{\pi}_2^k - C\left(\pi_i^k, \tilde{\pi}_i^k\right)$. We introduce the following convex

---

[9]Stöowhase and Traxler (2005) examined the effects of fiscal equalization policies.

and nonfiscally-deductible concealment cost $C\left(\pi_i^k, \tilde{\pi}_i^k\right)$, which is widely used in the literature; for example, see Haufler and Schjelderup (2000), Swenson (2001), Kind et al. (2005), Peralta et al. (2006), Devereux et al. (2008), Nielsen et al. (2008), and Keen and Konrad (2013):[10]

$$C\left(\pi_i^k, \tilde{\pi}_i^k\right) = 2\delta(e)\left(\pi_i^k - \tilde{\pi}_i^k\right)^2, \ i = 1, 2 \text{ and } k = a, b. \tag{1}$$

Several explanations are in order. First, $\delta(e)$ is a scaling factor for resource costs associated with profit shifting. It reflects the cost of hiring accounting experts to produce the required documents, expected penalties to be paid to the government, or the expected market sanction when caught cheating on tax liabilities. In the context of tax evasion, a standard assumption in the literature is that such costs are increasing and convex in the extent of profit shifting (tax evasion), $|\pi_i^k - \tilde{\pi}_i^k|$, regardless of the direction of profit shifting (i.e., it is cost equivalent to shift profits outward or inward).

Second, $\delta(e) = \delta(e_i, e_j)$ depends on the governments' enforcement efforts e_i, e_j, such as tougher monitoring, more efficient information sharing, and the efforts to negotiate and reach agreements with the other country's tax authority. $\delta(e)$ is an increasing function of e_i and e_j, such that stricter enforcement implies a higher $\delta(e)$. Moreover, in reality, dispersed (unilateral) enforcement efforts between involved countries are less effective in aggregate.[11] For instance, a lack of tax-relevant information provided by the host country makes the taxable income unclear to the home country, and the tax authorities cannot address tax fraud effectively. To formalize the imperfect substitutability of enforcement efforts, we adopt the following CES formula:

$$\delta(e_1, e_2) = \left(0.5 e_1^{-\rho} + 0.5 e_2^{-\rho}\right)^{-\frac{1}{\rho}}, \ \ \rho \geq -1. \tag{2}$$

The polar cases are: (i) $\rho = -1$ (perfect substitutes: total enforcement is based on the average enforcement); (ii) $\rho \to 0$ (the Cobb–Douglas case $\delta(e_1, e_2) = e_1^{0.5} e_2^{0.5}$); and (iii) $\rho \to \infty$ (the weakest–link case $\delta(e_1, e_2) = \min[e_1, e_2]$, where total enforcement is based on that of the lowest enforcer). For example, if during the mutual agreement procedure, the low-tax country can exercise a veto power on the transfer price and taxable incomes of the MNEs, then the enforcement technology becomes closer to the weakest–link formula.

---

[10]See also Huizinga and Laeven (2008) and Amerighi and Peralta (2010) for a slightly different specification. In a related model, Huizinga et al. (2008) considered MNEs that allocate tax allowances of debt finance (debt shifting), whereby allocation of the debt affects the expected bankruptcy costs, which are not tax deductible.

[11]Klassen and Laplante (2012) showed that profit shifting in a given country depends not only on the enforcement of the regulations in the home country but also on the implementation of the regulations in the host country.

The tax revenue in country i is:

$$R_i = t_i \left(\tilde{\pi}_i^a + \tilde{\pi}_i^b \right)$$

We assume that governments seek to maximize their fiscal revenue net of the enforcement cost (the tax administration costs). Adding the consumer surplus in the governmental objective function will not affect the analysis because firms' production and equilibrium prices are independent of tax and enforcement choices (see the proof of Lemma 1 for the derivation). This feature is similar to a widely used model by Kanbur and Keen (1993). We assume that $t_i \leq 1$, for $i = 1, 2$. Assuming a quadratic cost of enforcement $(c(e_i) = \eta \frac{(e_i)^2}{2})$ for simplicity, welfare in country i is:

$$W_i = R_i - \eta \frac{(e_i)^2}{2}, \tag{3}$$

where $\eta > 0$ is a parameter for the enforcement cost. To ensure an interior solution $(t_i \leq 1 \ (i = 1, \ 2))$ in the equilibrium, we assume that $\eta \geq 3$ (η is sufficiently high) throughout the rest of the paper.

2.2 Profit shifting by firms

The sequence of events is as follows. First, both countries set their enforcement efforts and tax rates. Subsequently, given enforcement and tax choices, multinational enterprises compete à la Cournot in each local market and choose a level of production in each country and the amount of profit to be shifted.

The issue of enforcement and tax timing will be discussed in the next subsection. In this subsection, we analyze the decisions of the firms in each country, given the tax $t = (t_1, t_2)$ and enforcement $e = (e_1, e_2)$ choices made earlier. Firm k $(k = a, \ b)$ chooses the quantities to produce in each market, (q_1^k, q_2^k) and the profit to report, $(\tilde{\pi}_1^k, \tilde{\pi}_2^k)$, to maximize the after-tax profit net of the profit-shifting cost, as follows:

$$(1 - t_1)\tilde{\pi}_1^k + (1 - t_2)\tilde{\pi}_2^k - 2\delta(e)(\pi_1^k - \tilde{\pi}_1^k)^2,$$

subject to $\tilde{\pi}_1^k + \tilde{\pi}_2^k = \pi_1^k + \pi_2^k$.

Recall that $\gamma_1 \geq \gamma_2$. Let $\epsilon \in [0, 1)$ be a parameter that represents the extent of asymmetry:

$$\epsilon \equiv \frac{1 - (\gamma_2/\gamma_1)^2}{1 + (\gamma_2/\gamma_1)^2},$$

(or $\dfrac{\gamma_1^2}{\gamma_1^2 + \gamma_2^2} = \dfrac{1+\epsilon}{2} \in [1/2, 1)$). We normalize $\beta = \dfrac{2}{9}(\gamma_1^2 + \gamma_2^2)$ so that the total profit $\sum_{k=a,b}(\pi_1^k + \pi_2^k)$ in Lemma 1 below is normalized to 1. Let $\tilde{\pi}_i = \tilde{\pi}_i^a + \tilde{\pi}_i^b$ be the total reported profit in country i.

Lemma 1 $\pi_1^k = \dfrac{\gamma_1^2}{9\beta} \equiv \dfrac{1+\epsilon}{4},\ \pi_2^k = \dfrac{1-\epsilon}{4}\ (k = a,\ b),\ \tilde{\pi}_1^a + \tilde{\pi}_1^b = \dfrac{1+\epsilon}{2} -$

$\dfrac{t_1 - t_2}{2\delta(e)} \equiv \tilde{\pi}_1(t,e),\ \tilde{\pi}_2^a + \tilde{\pi}_2^b = \dfrac{1-\epsilon}{2} + \dfrac{t_1 - t_2}{2\delta(e)} \equiv \tilde{\pi}_2(t,e).$

Proof: See the Appendix.

The reported profit in country i ($\tilde{\pi}_i(t,e)$) consists of actual profits $\pi_i = \pi_i^a + \pi_i^b$ and the amount of profit shifting $\pi_i - \tilde{\pi}_i$. The former solely depends on the market size parameter ϵ. The latter is proportional to the tax difference $t_i - t_j$ and inversely proportional to the total enforcement $\delta(e)$. The tax base elasticity (the sensitivity of the tax base to fiscal rates) is higher in the small country,[12] which is in line with existing asymmetric tax competition models.

2.3 Enforcement and tax timing

From Lemma 1, we have $R_i(t,e) = t_i \tilde{\pi}_i(t,e)$. The net tax revenue (3), taking account of the firms' behavior, is:

$$W_i = t_i \tilde{\pi}_i(t_i, t_j, e) - \eta \frac{(e_i)^2}{2} = t_i \left(\frac{1+\epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)} \right) - \eta \frac{(e_i)^2}{2}, \qquad (4)$$

where $\epsilon_1 = \epsilon = -\epsilon_2$.

First, we describe the case where the tax t_i and the enforcement e_i are chosen simultaneously and noncooperatively. In the Appendix, we show the following:

Lemma 2 *Suppose that t_i and e_i ($i = 1, 2$) are chosen simultaneously and noncooperatively. For any ρ and ϵ, there is no equilibrium with positive taxes in pure strategy.*

Proof: See the Appendix.

The blue surface in Figure 1 illustrates country 1's welfare with respect to (e_1, t_1), against $(e_2, t_2) = (e_2^n, t_2^n)$, which satisfies the first-order conditions

---

[12]The tax base elasticities at $t_1 = t_2 = t$ are $-\dfrac{\partial \tilde{\pi}_1}{\partial t_1} \dfrac{t_1}{\tilde{\pi}_1} = \dfrac{t}{\delta(e)(1+\epsilon)}$ and $-\dfrac{\partial \tilde{\pi}_2}{\partial t_2} \dfrac{t_2}{\tilde{\pi}_2} = \dfrac{t}{\delta(e)(1-\epsilon)}$.
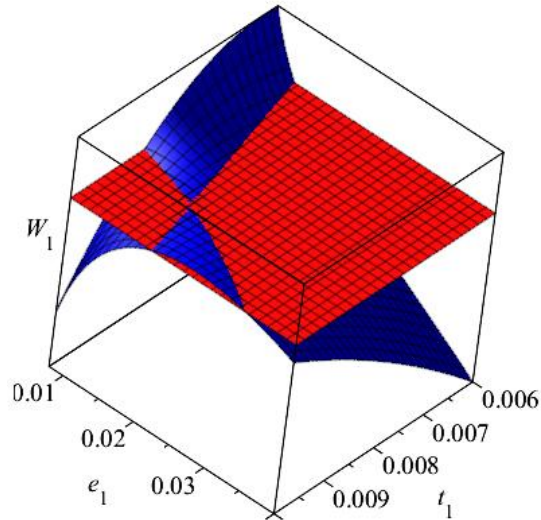
Figure 1: The blue surface indicates $W_1(e_1, t_1, e_2^n, t_2^n)$ for $\rho = -1$ and $\epsilon = 0.1$. The red panel indicates the welfare level corresponding to $W_1 = W_1(e_1^n, t_1^n, e_2^n, t_2^n)$.

of country 2. We set $\rho = -1$ and $\epsilon = 0.1$. The red panel indicates the welfare level corresponding to the value of $(e_1, t_1) = (e_1^n, t_1^n)$ that uniquely satisfies the local first-order conditions; it is a saddle point, as there is scope for welfare improvement for country 1 by jointly decreasing (or increasing) its tax and enforcement levels. For all $\rho < 0$, this saddle point property is inherent in the present model: there is no pure strategy equilibrium.

Another feature of this setting is that the low-tax country has no incentive for enforcement ($e_2^n = 0$). For $\rho \geq 0$ in (2), i.e., when there is stronger complementarity, including the Cobb–Douglas and weakest–link cases, this means that $\delta(e_1, 0) = 0$ for all e_1 (both enforcement contributions are indispensable for $\delta(e) > 0$). Hence, from (1), the firms engage in unlimited profit shifting so that the countries cannot tax profits.

For the rest of this paper, we adopt the following sequence of enforcement and tax decisions: first, both countries set their enforcement efforts, and second, both countries choose their tax rates. The reasons that we adopt this setup are as follows. First, the level of enforcement effort is determined by specific rules and laws of monitoring, inspection and information sharing, which are less reversible in nature than the tax rates, which can be changed

9

more easily (see Bacchetta and Espinosa (1995), Peralta et al. (2006), Bucovetsky and Haufler (2008), and Keen and Konrad (2013)).[13] Second, by treating tax-enforcement decisions as long-term policy variables, this structure allows us to examine a current issue regarding the possibility of enforcement cooperation under international tax competition.

3 Equilibrium choices and enforcement complementarity

3.1 Tax choices

Given the sequence of decisions, where enforcement efforts are chosen first and taxes are chosen second (followed by the MNEs' decisions characterized in Lemma 1), the model is solved by backward induction. In this subsection, we first derive the governments' choices of taxes.

Given the enforcement level $e = (e_i, e_j)$, each country noncooperatively chooses its own tax rate t_i $(i = 1, 2)$ to maximize the welfare in (4). The first-order conditions are:

$$\frac{\partial W_i}{\partial t_i} = \frac{1 + \epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)} + t_i \frac{-1}{2\delta(e)} = 0, \tag{5}$$

where $\epsilon_1 = \epsilon = -\epsilon_2$. The second-order conditions are satisfied. They yield the following equilibrium taxes, denoted by $(t_1^N(e), t_2^N(e))$:

$$t_1^N(e) = \delta(e)\left(\frac{3 + \epsilon}{3}\right) \quad \text{and} \quad t_2^N(e) = \delta(e)\left(\frac{3 - \epsilon}{3}\right). \tag{6}$$

Reflecting its greater market size, the large country imposes higher taxes, and the tax rate differential widens with greater asymmetry: $t_1^N(e) - t_2^N(e) = 2\frac{\delta(e)\epsilon}{3} > 0$. From Lemma 1, both firms shift profits from the large to the small (low-tax) country: for $\tilde{\pi}_i^N \equiv \tilde{\pi}_i(t_1^N(e), t_2^N(e), e)$, we have $\tilde{\pi}_1^N = \frac{1 + \epsilon}{2} - \frac{t_1^N(e) - t_2^N(e)}{2\delta(e)} = \frac{3 + \epsilon}{6}$ and $\tilde{\pi}_2^N = \frac{3 - \epsilon}{6}$. The tax revenues $R_i^N(e) = t_i^N(e)\tilde{\pi}_i^N$ $(i = 1, 2)$ are as follows:

$$R_1^N(e) = \frac{\delta(e)}{2}\left(\frac{3 + \epsilon}{3}\right)^2 \geq R_2^N(e) = \frac{\delta(e)}{2}\left(\frac{3 - \epsilon}{3}\right)^2 \quad \text{for all } \epsilon \geq 0. \tag{7}$$

Note that, given e, the total revenue $R_1^N(e) + R_2^N(e)$ is increasing as the extent of country asymmetry increases.

---

[13]In Section 3.2, we briefly discuss the case where taxes are chosen first and the enforcement level is chosen later.

3.2 Noncooperative enforcement choices

In the first stage, the governments in each country choose their enforcement effort levels, taking into account the behavior in the subsequent stages. In this subsection, we first examine noncooperative enforcement choices, where each country chooses e_i $(i = 1, 2)$ simultaneously and independently.

Given e_j, country i maximizes $W_i^N(e_i, e_j) = R_i^N(e_i, e_j) - \eta \dfrac{(e_i)^2}{2}$ where $R_i^N(e)$'s are given in (7) and $\delta(e_1, e_2)$ takes the form of (2). The first-order conditions are given by:[14]

$$\frac{\partial \delta(e)}{\partial e_i} \frac{1}{2} \left(\frac{3 + \epsilon_i}{3} \right)^2 - \eta e_i = 0, \tag{8}$$

with $\epsilon_1 = \epsilon = -\epsilon_2$.

Let $(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon))$ be the enforcement level at the noncooperative equilibrium with ρ and ϵ given. From (8), we obtain the following:

Lemma 3 $e_1^N(\rho, \epsilon) > e_2^N(\rho, \epsilon) > 0$ *for all $\rho < \infty$ and $\epsilon > 0$.*

Proof: See the Appendix.

Lemma 3 illustrates the synergy between market size and enforcement. From (7), the marginal benefit of enforcement is proportional to the tax revenue. Under asymmetry, country 1 exhibits higher equilibrium taxes. In turn, the country is playing for high stakes in the enforcement game, which leads to the higher equilibrium level of enforcement. This feature corresponds to the real world situation, where the enforcement of legislation differs significantly across countries, with low-tax countries typically having weaker enforcement regimes.[15]

In the Appendix, we show the following:

---

[14]A general treatment is as follows:

$$\frac{\partial W_i^N}{\partial e_i} = \frac{\partial R_i(t_i, t_j, e)}{\partial e_i} + \frac{\partial R_i(t_i, t_j, e)}{\partial t_j} \frac{\partial t_j^N}{\partial e_i} - c'(e_i) = 0.$$

The first term is negative for the low-tax country and positive for the high-tax country, and the second term is positive for both countries. Enforcement efforts increase the cost of profit shifting, which in turn induces higher taxes.

[15]When the sequence of decisions is reversed, so that taxes are chosen first and the enforcement level is chosen later, then: (i) a subgame-perfect equilibrium with positive taxes exists only if $\rho < 0$ (i.e., if enforcement efforts are substitutes) and $\epsilon \geq 0.2$, and (ii) both the equilibrium taxes and enforcement efforts are lower than those obtained in (6) and Lemma 3 (this derivation is available upon request to the authors). In the equilibrium of this reverse timing scenario, the enforcement effort is undertaken only by country 1

Lemma 4 *(i)* $e_1^N(\rho, \epsilon)$ *is decreasing in* ρ *(enforcement complementarity) and* $e_2^N(\rho, \epsilon)$ *is increasing in* ρ *for all* ρ *and* $\epsilon > 0$.

(ii) $e_2^N(\rho, \epsilon)$ *is decreasing in* ϵ *and* $e_1^N(\rho, \epsilon) - e_2^N(\rho, \epsilon)$ *is increasing in* ϵ *for all* ρ *and* ϵ. *The effect of increasing* ϵ *on* $e_1^N(\rho, \epsilon)$ *is ambiguous.*

(iii) $\delta^N(\rho, \epsilon) \equiv \delta(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon))$ *is decreasing in* ρ *for all* ρ *and* $\epsilon > 0$. $\delta^N(\rho, \epsilon)$ *is increasing in* ϵ *when* $\rho < -\dfrac{2}{3}$ *and* $\delta^N(\rho, \epsilon)$ *is decreasing in* ϵ *when* $\rho > -\dfrac{2}{3}$.
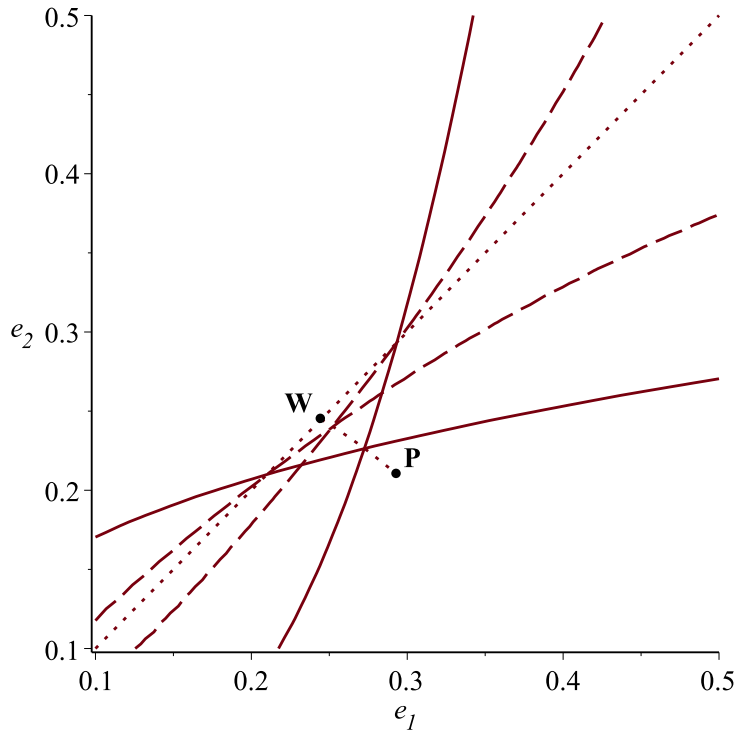
Proof: See the Appendix.

Figure 2: Enforcement reaction functions for $\rho = -0.2$, $\epsilon = 0.25$ (solid curves) and for $\rho = 5$, $\epsilon = 0.25$ (dashed curves). The curve **PW** is the locus of $(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon))$ for $\epsilon = 0.25$.

($e_1 > 0$ and $e_2 = 0$). Although e_1 is shown to be increasing in the tax gap $t_1 - t_2$, this means that the low-tax country will have an incentive to increase its tax t_2 to lower the enforcement of the other country (e_1). In contrast, in the present analysis, where taxes are chosen after the enforcement, increases in enforcement efforts induce higher taxes because taxes are positively related to $\delta(e)$ in (6).

Figure 2 illustrates the enforcement reaction function of country i, derived from (8), for the cases of $\rho = -0.2$, $\epsilon = 0.25$ and $\rho = 5$, $\epsilon = 0.25$. The reaction functions are upward-sloping, i.e., there are strategic complementarities in enforcement efforts.[16] The dotted curve **PW** illustrates the locus of the equilibrium enforcement efforts (the intersection of the reaction functions) for $\epsilon = 0.25$. From $\mathbf{P} = (e_1^N(-1, \epsilon), e_2^N(-1, \epsilon))$, $e_1^N(\rho, \epsilon)$ decreases in ρ, whereas $e_2^N(\rho, \epsilon)$ increases in ρ, converging to $\mathbf{W} = (e_1^N(\infty, \epsilon), e_2^N(\infty, \epsilon))$ where $e_1^N(\infty, \epsilon) = e_2^N(\infty, \epsilon)$.
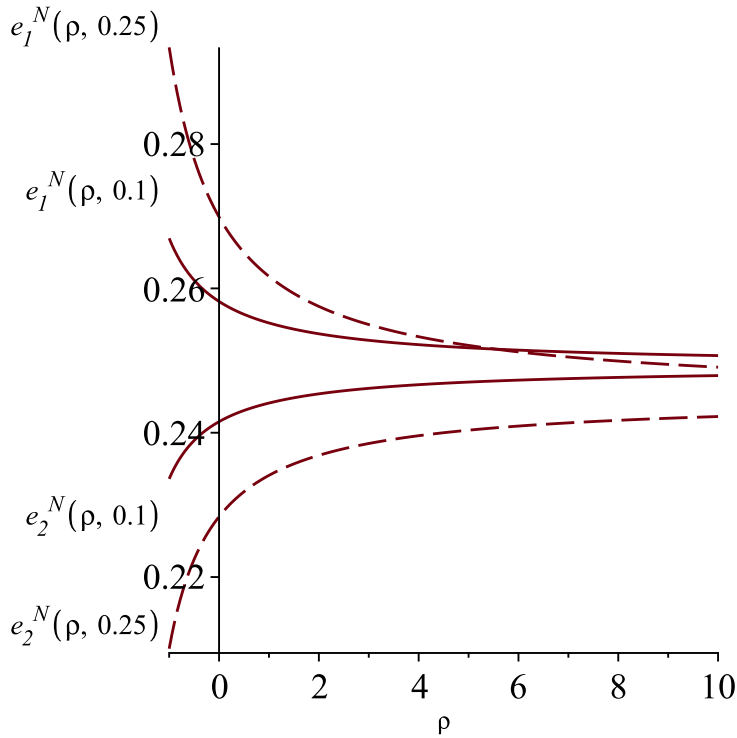

Figure 3: Noncooperative enforcement level for $\epsilon = 0.1$ and $\epsilon = 0.25$

Figure 3 illustrates $e_1^N(\rho, \epsilon)$ and $e_2^N(\rho, \epsilon)$ for different values of ϵ. It shows that $e_2^N(\rho, 0.25)$ is lower than $e_2^N(\rho, 0.1)$ for all ρ. Regarding $e_1^N(\rho, \epsilon)$, $e_1^N(\rho, 0.25)$ may be higher or lower than $e_1^N(\rho, 0.1)$, depending on the value of ρ.[17] In aggregate, stronger complementarity gives more power to the low-tax country in setting the noncooperative enforcement level $\delta^N(\rho, \epsilon)$. $\delta^N(\rho, \epsilon)$ is

---

[16]For $e_2 < e_1$, greater complementarity (larger ρ) reduces the best-response effort of country 1 and increases that of country 2.

[17]An increase in country 1's market size ϵ tends to increase $e_1^N(\rho, \epsilon)$. However, it also

13

increasing in ϵ when ρ is sufficiently low, reflecting the facts that equilibrium enforcement is proportional to the tax revenue and that total revenue is increasing in ϵ. However, when ρ is high, enforcement becomes more constrained by the low-tax country and $\delta^N(\rho, \epsilon)$ becomes decreasing in ϵ.

3.3 Benefit of enforcement cooperation

Now, we examine cooperative enforcement choices in the first stage and see whether the cooperative framework is adopted unanimously. Here, both countries choose their enforcement levels to maximize their joint welfare. This reflects an agreement regarding the level of information exchange in the tax treaty. However, in a reflection of the fact that most tax treaties do not have cooperative or harmonized tax rates in the real world, we assume that countries set taxes (t_i) noncooperatively. Therefore, countries choose $e = (e_1, e_2)$, anticipating the noncooperative tax game $(t_1^N(e), t_2^N(e))$ and tax revenues $(R_1^N(e), R_2^N(e))$ in (7). That is:

$$\max_{e_1, e_2} \sum_i \left(R_i^N(e_i, e_j) - \eta \frac{(e_i)^2}{2} \right). \tag{9}$$

We obtain $e_1^{N*}(\rho, \epsilon) = e_2^{N*}(\rho, \epsilon) = \dfrac{9 + \epsilon^2}{18\eta}$, which is invariant with respect to ρ (see the proof of Proposition 1). Enforcement efficiency requires both countries to exert the same enforcement efforts because of the convex cost function. Compared with the noncooperative solution, enforcement cooperation *doubles* the total level of enforcement. This is because the positive fiscal externality of enforcement $\partial R_j^N / \partial e_i > 0$ for $i \neq j$ is now internalized.

Let $W_i^N(e^N(\rho, \epsilon))$ and $W_i^N(e^{N*}(\rho, \epsilon))$ $(i = 1, 2)$ be the welfare levels in the noncooperative and cooperative regimes, respectively. The large country unambiguously gains by enforcement cooperation $(W_1^N(e^{N*}(\rho, \epsilon)) > W_1^N(e^N(\rho, \epsilon))$ for all ρ and ϵ), owing to better coordination in enforcement efforts. However, for country 2, there are offsetting effects. As $e_2^{N*}(\rho, \epsilon)$ increases in ϵ but $e_2^N(\rho, \epsilon)$ decreases in ϵ (Lemma 4), the cost of coordinated enforcement increases in ϵ. As a result, even though the tax revenue increases $(R_2^N(e^{N*}(\rho, \epsilon)) > R_2^N(e^N(\rho, \epsilon)))$, country 2 may prefer the noncooperative regime. This feature corresponds to a small country being unwilling to cooperate when tax rates are the only instrument of tax competition (Bucovetsky (1991) and Wilson (1991)). However, we go further than this. When the relevant parameter to be coordinated is tax enforcement, the following

---

decreases $e_2^N(\rho, \epsilon)$. When ρ is sufficiently high, the latter tends to decrease $e_1^N(\rho, \epsilon)$ owing to strategic complementarity.

proposition shows the viability of cooperation with respect to the levels of asymmetry (ϵ) and enforcement complementarity (ρ):

Proposition 1 *(i) For the high-tax country, $W_1^N(e_1^{N*}(\rho, \epsilon), e_2^{N*}(\rho, \epsilon)) > W_1^N(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon))$ for all parameter values. For the low-tax country, there exists a cutoff level of asymmetry $\epsilon^N(\rho)$ such that $W_2^N(e_1^{N*}(\rho, \epsilon), e_2^{N*}(\rho, \epsilon)) > W_2^N(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon))$ iff $\epsilon < \epsilon^N(\rho)$.*
    (ii) $\epsilon^N(\rho)$ is an increasing function of ρ and belongs to the interval $(0.25,\ 0.41)$.
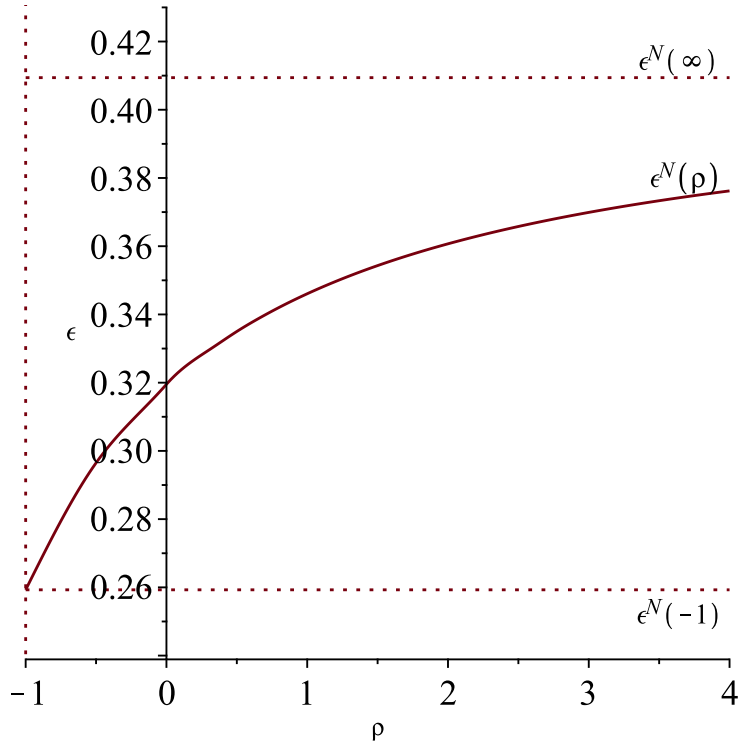
Proof: See the Appendix.

Figure 4: Critical value of asymmetry $\epsilon^N(\rho)$

For a given level of enforcement complementarity (ρ), the gap between $e_1^N(\rho, \epsilon)$ and $e_2^N(\rho, \epsilon)$ increases when ϵ increases. Owing to the conflict of interest with respect to the desired level of the enforcement, country 2 prefers the noncooperative regime when asymmetry is above a certain cutoff value. However, this cutoff level of asymmetry is increasing in ρ, suggesting that the low-tax country is more likely to benefit from cooperation with greater

15

enforcement complementarity. In the (ρ, ϵ) space, Figure 4 indicates where cooperation is mutually beneficial (below/right of the curve $\epsilon^N(\rho)$).

Proposition 1 suggests several cases where the enforcement cooperation (by unanimity) is more or less likely. Consider, for example, a situation where the lack of tax-relevant information by the host (low-tax) country lowers the overall enforcement level ($\delta^N(\rho, \epsilon)$). This in itself encourages the profit shifting that seems to favor the low-tax country. However, such a compliance structure negatively affects subsequent tax choices and the equilibrium tax revenue in (7). In such cases, the low-tax country may agree on a cooperative enforcement level that enables better inspection and information sharing. In contrast, if overall tax enforcement is additively separable with respect to the transfer pricing documentation from each country or the number of tax officials in each country (indicating high substitutability in enforcement), then the low-tax country may prefer to forego enforcement cooperation: in such cases, the free-riding by the low-tax, low-enforcement country still allows for a relatively high $\delta^N(\rho, \epsilon)$ in a noncooperative equilibrium.

4 Follow the leader

In this section, we consider a second driver of enforcement cooperation: the case of tax leadership. In this case, the high-tax country as the tax leader would commit to a tax rate at some level, which gives a tax-follower benefit to the low-tax country. We shall see how the tax leadership motivates the low-tax country to agree on cooperative enforcement efforts. The argument for the form of leadership by the high-tax country (the large country) follows Hindriks and Nishimura (2015, 2017).[18]

We continue to consider that the enforcement level is set before taxes are set, and we solve the game by backward induction. Given the enforcement choices $e = (e_1, e_2)$, the (small) country 2, as the tax follower, chooses t_2, given t_1. Country 1, as the Stackelberg leader, maximizes $W_1(t_1, t_2, e)$ with respect to t_1 along country 2's tax reaction function $t_2 = \hat{t}_2(t_1; e)$. This interaction yields the equilibrium tax rates denoted by $(t_1^S(e), t_2^S(e))$ and the tax revenues $R_i^S(e) = t_i^S(e)\tilde{\pi}_i(t_1^S(e), t_2^S(e), e)$ $(i = 1, 2)$ with the following properties:

Lemma 5 *For any given level of enforcement $e = (e_1, e_2)$ we have:*
(i) $t_i^S(e) > t_i^N(e)$ $(i = 1, 2)$.
(ii) $R_i^S(e) > R_i^N(e)$ $(i = 1, 2)$ and $t_1^S(e) - t_2^S(e) > t_1^N(e) - t_2^N(e) > 0$.

---

[18] At the end of this section, we briefly discuss the case of the small country's leadership.

(iii) There exist values of $\hat{\epsilon}$ and $\tilde{\epsilon}$, $\hat{\epsilon} > \tilde{\epsilon}$, such that $R_1^S(e) \gtrless R_2^S(e) \iff \epsilon \gtrless \hat{\epsilon}$, and $|R_1^S(e) - R_2^S(e)| \lessgtr R_1^N(e) - R_1^N(e) \iff \epsilon \gtrless \tilde{\epsilon}$.

Proof: See the Appendix.

We have $\partial \hat{t}_2(t_1; e)/\partial t_1 \in (0, 1)$, as in conventional models. Owing to the strategic complementarities in taxes, equilibrium taxes are higher ($t_i^S(e) > t_i^N(e)$ ($i = 1, 2$) in part (i)). Since the slopes of the tax response functions are less than one, the tax increase by the tax follower is less than that of the tax leader (i.e., the widening of the tax gap referred to in part (ii) occurs). Regarding part (iii), when ϵ (asymmetry) is small enough, the tax leader collects less tax revenue than the tax follower, and vice versa when ϵ is large enough.[19] This feature contrasts with the Nash tax competition where the tax-revenue difference $R_1^N(e^N) - R_2^N(e^N) > 0$ is increasing in ϵ. As a result, the dispersion of tax revenues is smaller under Stackelberg competition than under Nash tax competition, if ϵ is sufficiently large.

Then, we move backward to compute the noncooperative enforcement choices, followed by the Stackelberg taxation. Given e_j, country i maximizes $W_i^S(e_i, e_j) = R_i^S(e_i, e_j) - \eta \dfrac{(e_i)^2}{2}$. Let $e^S = (e_1^S, e_2^S)$ be the noncooperative equilibrium enforcement. In this section, we restrict our attention to the case of perfect substitutability of enforcement efforts ($\rho = -1$):

$$\delta(e_1, e_2) = 0.5 \, e_1 + 0.5 \, e_2 \tag{10}$$

The aggregate enforcement level $\delta(e)$ positively affects equilibrium tax rates. The question is how the enforcement efforts will be changed by the tax leadership. Denoting $e_i^N(-1, \epsilon) = e_i^N$ ($i = 1, 2$) and $\delta^N(-1, \epsilon) = \delta^N$ with a given ϵ, we show the following:

Lemma 6 *(i) $e_1^S > e_1^N$, $e_2^S > e_2^N$, $\delta(e_1^S, e_2^S) \equiv \delta^S > \delta^N$. That is, under the Stackelberg tax competition, both the leader and the tax follower exert more effort than under the Nash tax game.*
(ii) $R_i^S(e^S) > R_i^N(e^N)$ ($i = 1, 2$) and $t_1^S(e^S) - t_2^S(e^S) > t_1^N(e^N) - t_2^N(e^N)$.
(iii) $|e_1^S - e_2^S| \lessgtr e_1^N - e_2^N \iff \epsilon \gtrless \tilde{\epsilon}$. That is, the Stackelberg tax competition reduces the dispersion of the enforcement efforts, when asymmetry is sufficiently large.

Proof: See the Appendix.

---

[19]In the Appendix, we show that the reported tax base is smaller for the tax leader ($\tilde{\pi}_1(t_1^S(e), t_2^S(e), e) < \tilde{\pi}_2(t_1^S(e), t_2^S(e), e)$). In combination with $t_1^S(e) > t_2^S(e)$, this has an offsetting effect on the tax revenue $R_i^S(e)$ ($i = 1, 2$).

From $R_i^S(e) > R_i^N(e)$ $(i = 1, 2)$ in Lemma 5.(ii), the marginal benefit of enforcement is increased for both countries under the tax leadership, so that $e_i^S > e_i^N$ $(i = 1, 2)$ in Lemma 6.(i). As for part (ii), similarly to the case of a "minimum tax" over the Nash tax competition in Kanbur and Keen (1993), the benefit for the high-tax country arises from the higher tax revenue along the tax reaction function $(R_1^S(e^S) > R_1^N(e^N))$. At the same time, the low-tax country benefits from the profit shifting.[20] Here, the wider tax differential $(t_1^S(e^S) - t_2^S(e^S) > t_1^N(e^N) - t_2^N(e^N))$ induces more profit shifting. As emphasized in Kanbur and Keen (1993), this type of coordination strategy is different from tax harmonization, which eliminates the benefit of the profit shifting to the low-tax country. These features lead to Lemma 6.(iii) regarding the dispersion of the enforcement efforts.

Next, we examine the cooperative solution and compare it with the non-cooperative enforcement decisions, where the tax choices take place under the Stackelberg competition. As in Section 3.3, both countries choose their enforcement levels $e = (e_1, e_2)$ to maximize their joint welfare, but the countries subsequently compete in taxes. Therefore, given e, countries choose $(t_1^S(e), t_2^S(e))$. The choice of e_i and e_j is given by $\max_{e_1, e_2} \sum_i \left(R_i^S(e_i, e_j) - \eta \frac{(e_i)^2}{2} \right)$, yielding $e = (e_1^{S*}, e_2^{S*})$. Under the Nash game (with $\rho = -1$) in Proposition 1 (Figure 4), the low-tax country prefers the noncooperative regime $(W_2^N(e_1^{N*}(-1, \epsilon), e_2^{N*}(-1, \epsilon)) < W_2^N(e_1^N, e_2^N))$ if and only if $\epsilon > \epsilon^N(-1)$. In comparison to that case, we show the following:

Proposition 2 *Cooperation on enforcement is more likely under the Stackelberg tax competition than under the Nash tax competition: $W_1^S(e_1^{S*}, e_2^{S*}) > W_1^S(e_1^S, e_2^S)$ for all ϵ, whereas $W_2^S(e_1^{S*}, e_2^{S*}) > W_2^S(e_1^S, e_2^S)$ $(i = 1, 2)$ iff $\epsilon < \epsilon^S$, where ϵ^S is greater than $\epsilon^N(-1)$ in Proposition 1.*

Proof: See the Appendix.

The tax leadership (by the high-tax country) facilitates the participation of the low-tax country in the enforcement cooperation. There are two reasons for this result. First, tax revenues are higher under tax leadership $(R_i^S(e^S) > R_i^N(e^N)$ $(i = 1, 2))$, as in Hindriks and Nishimura (2015), so that the benefit of internalizing the enforcement externality is larger. Second, the enforcement gap is reduced, as shown in Lemma 6.(iii), which reduces the

---

[20]The difference between the minimum tax case and the case of leadership by the high-tax country is that the former entails a move along the high-tax country's reaction function, whereas here, in the latter case, we consider a move along the low-tax country's reaction function.

extent of the disagreement with respect to the desired level of enforcement. This result continues to hold for all ρ in (2).[21]

Finally, we discuss the issue of leadership.[22] Let $R_i^M(e)$ $(i = 1, 2)$ be country i's revenue under country 2's leadership, with a given level of e. Then, we have $R_1^M(e) > R_1^S(e) > R_1^N(e)$ and $R_2^S(e) > R_2^M(e) > R_2^N(e)$. That is, country 2's leadership also increases the countries' enforcement incentives compared with the Nash tax competition. However, regarding the equilibrium enforcement effort e_i^M $(i = 1, 2)$, we have $e_1^M - e_2^M > e_1^N - e_2^N$ for all ϵ (the enforcement gap is increased with the reverse leadership). As a result, the cutoff value of asymmetry sustaining cooperation is even smaller than that in Proposition 1. Moreover, we can show that the high-tax leadership equilibrium Pareto dominates the low-tax leadership equilibrium under sufficient asymmetry.[23]

5 Conclusion

We developed an asymmetric tax-competition model with two policy instruments, the corporate tax rate and the enforcement level. We characterized the model with simple closed-form solutions that facilitate analyzing the determinants of voluntary cooperation for tax enforcement. Stronger complementarity gives more power to the low-tax country in setting a noncooperative enforcement level. In reality, the low-tax country country may be able to exercise a veto power towards tighter enforcement during, for example, a mutual agreement procedure on the taxable incomes of the MNEs. Surprisingly, we showed that enforcement cooperation is most likely to occur when the chain of enforcement is as strong as its weakest–link.

In the Stackelberg tax competition where the large country leads, the equilibrium taxes become higher and the profit shifting *increases*. As a result, both the tax leader and the tax follower will exert greater enforcement efforts than under the simultaneous tax choice scenario. Thus, enforcement

---

[21]For all $\rho > -1$, we can show that the critical value for cooperation $\epsilon^S(\rho)$ is higher than $\epsilon^N(\rho)$ of Proposition 1.

[22]A brief review of the background literature is as follows. Following Kanbur and Keen's (1993) cross-border shopping model, Wang (1999) assumed that the high-tax country behaves as a Stackelberg leader, and showed that both countries become better off because of the tax leadership. In the present model, where asymmetry is defined by the market size, when δ is fixed, the high-tax country's leadership results in the equilibrium timing à la Hamilton and Slutsky (1990). See Hindriks and Nishimura (2015, p.68).

[23]For the cooperative effort level (e_1^{M*}, e_2^{M*}), we have $W_2^M(e_1^{M*}, e_2^{M*}) < W_2^M(e_1^M, e_2^M)$ for all $\epsilon > \epsilon^N(-1)$. Also, $W_2^M(e_1^M, e_2^M) < W_2^S(e_1^S, e_2^S)$ for all ϵ and $W_1^M(e_1^M, e_2^M) < W_1^S(e_1^S, e_2^S)$ if ϵ is sufficiently high.

cooperation will be more likely under the Stackelberg tax competition. Interestingly, this type of coordination strategy is different from tax harmonization, which eliminates the benefit of profit shifting to the small country.

The action programs of the OECD BEPS project have various degree of commitment, ranging from strong commitment to the consistent implementation of the program across countries, to the weak form of commitment, where countries are free to assess and pick only the policies that they wish to implement. This paper has made two contributions to the literature. First, by developing a simple model that captures the central features of profit shifting, we have provided significant insights into the costs and benefits of enforcement cooperation. Second, the analysis provides guidance on what empirical quantities to seek in determining the viability of enforcement cooperation in practice.

Appendix

Proof of Lemma 1: The optimization problem of firm k $(k = a, b)$ is equivalent to:

$$\max_{q_1^k, q_2^k, \tilde{\pi}_1^k} \quad (1 - t_1)\tilde{\pi}_1^k + (1 - t_2)[\pi_1^k + \pi_2^k - \tilde{\pi}_1^k] - 2\delta(e)(\pi_1^k - \tilde{\pi}_1^k)^2,$$

subject to $\pi_i^k = \{\gamma_i - \beta(q_i^k + q_i^{k'})\}q_i^k$, $i = 1, 2$, $k' \neq k$.

The first-order condition for $\tilde{\pi}_1^k$ yields:

$$-t_1 + t_2 - 4\delta(e)(\tilde{\pi}_1^k - \pi_1^k) = 0. \tag{11}$$

Regarding the choice of q_1^k and q_2^k, we have:

$$(1 - t_2 + 4\delta(e)(\tilde{\pi}_1^k - \pi_1^k))\frac{\partial \pi_1^k}{q_1^k} = (1 - t_1)[\gamma_1 - 2\beta q_1^k - \beta q_1^{k'}] = 0,$$

$$(1 - t_2)[\gamma_2 - 2\beta q_2^k - \beta q_2^{k'}] = 0,$$

where we use (11). These first-order conditions yield the reaction function $q_i^k = (\gamma_i - \beta q_i^{k'})/(2\beta)$, from which we have $q_i^a = q_i^b = \gamma_i/(3\beta)$ and $p_i = \gamma_i/3 + c_i$ $(i = 1, 2)$, which are independent of taxes and enforcement levels. In addition, we have $\pi_i^a = \pi_i^b = \gamma_i^2/(9\beta)$ and, from the normalization of β and the definition of ϵ in the text, we have $\sum_{k=a,b}(\pi_1^k + \pi_2^k) = 1$, $\pi_1^k = \dfrac{1 + \epsilon}{4}$ and $\pi_2^k = \dfrac{1 - \epsilon}{4}$ $(k = a, b)$. From (11), we obtain: $\tilde{\pi}_1^k = \pi_1^k - \dfrac{t_1 - t_2}{4\delta(e)}$ and

20

$\tilde{\pi}_2^k = \pi_2^k + \pi_1^k - \tilde{\pi}_1^k = \pi_2^k + \dfrac{t_1 - t_2}{4\delta(e)}$ ($k = a, b$). Therefore, the conclusion of the lemma holds. *Q.E.D.*

Proof of Lemma 2: When t_i and e_i are chosen simultaneously and independently, the first-order conditions of the welfare maximization with respect to t_i and e_i are:

$$\frac{\partial W_i}{\partial t_i} = \frac{1 + \epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)} + t_i \frac{-1}{2\delta(e)} = 0, \tag{12}$$

$$\frac{\partial W_i}{\partial e_i} = t_i \left(\frac{t_i - t_j}{2(\delta(e))^2} \right) \frac{\partial \delta(e_i, e_j)}{\partial e_i} - \eta e_i \leq 0. \tag{13}$$

First, consider the case of $\rho < 0$. (12) derives the preferred tax level of country i, where $t_1 > t_2$ for $\epsilon \in (0, 1)$ and $t_1 = t_2$ for $\epsilon = 0$. Taking these into account and given that $\partial \delta(e_i, e_j)/\partial e_i > 0$, (13) implies that country 2 chooses $e_2 = 0$. From (12) and (13), the set of values that satisfy the first-order necessary conditions is uniquely given by $e_1^n = \dfrac{\epsilon(3 + \epsilon)}{9\eta} 0.5^{-\frac{1}{\rho}}$, $e_2^n = 0$ and $t_1^n = 0.5^{-\frac{1}{\rho}} e_1^n \left(\dfrac{3 + \epsilon}{3} \right)$ and $t_2^n = 0.5^{-\frac{1}{\rho}} e_1^n \left(\dfrac{3 - \epsilon}{3} \right)$. Figure 1 shows $W_1(e_1, t_1, e_2^n, t_2^n)$ for the case of $\rho = -1$ and $\epsilon = 0.1$. For $e_1^a = \dfrac{\epsilon(3 + \epsilon)}{72\eta} 0.5^{-\frac{1}{\rho}} < e_1^n$ and $t_1^a = \dfrac{\epsilon(27 - 5\epsilon)(3 + \epsilon)}{432\eta} 0.5^{-\frac{2}{\rho}} < t_1^n$, we have $W_1(e_1^a, t_1^a, e_2^n, t_2^n) - W_1(e_1^n, t_1^n, e_2^n, t_2^n) = \dfrac{49\epsilon(3 + \epsilon)(\epsilon^2 - 15\epsilon + 18)0.5^{-\frac{2}{\rho}}}{10368\eta}$, which is positive for $\epsilon > 0$. For $\epsilon = 0$, we set $e_1^b = \dfrac{1}{18\eta} 0.5^{-\frac{1}{\rho}}$ and $t_1^b = \dfrac{1}{36\eta} 0.5^{-\frac{2}{\rho}}$ ($e_1^b > e_1^n = 0$ and $t_1^b > t_1^n = 0$), and we have $W_1(e_1^b, t_1^b, e_2^n, t_2^n) - W_1(e_1^n, t_1^n, e_2^n, t_2^n) = \dfrac{7}{1296\eta} 0.5^{-\frac{2}{\rho}} > 0.$[24] Therefore, the equilibrium does not exist.

Next, consider the case of $\rho \geq 0$. From (13), one must have $e_1 = 0$ or $e_2 = 0$. Note that, if $e_i = 0$ ($i = 1$ or 2) and $\rho \geq 0$ in (2), then $\delta(0, e_j) = 0$. Therefore, there is no equilibrium with $\delta(e) > 0$. Hence, from (1), the firms engage in unlimited profit shifting so that the equilibrium taxes cannot be positive. *Q.E.D.*

Proof of Lemma 3 and Lemma 4: (8) is equivalent to

$$e_i^{-\rho-1} \left(0.5 e_1^{-\rho} + 0.5 e_2^{-\rho} \right)^{\frac{1+\rho}{-\rho}} \left(\frac{3 + \epsilon_i}{6} \right)^2 - \eta e_i = 0 \ (i = 1, 2). \text{ Solving the sys-}$$

---

[24] $\lim_{\epsilon \to 0} W_1(e_1^n, t_1^n, e_2^n, t_2^n) = 0.$

tem, we have $e_1^N(\rho, \epsilon) = \dfrac{1}{\eta} \left(\dfrac{3+\epsilon}{6} \right)^2 \left(0.5 \left(\dfrac{3-\epsilon}{3+\epsilon} \right)^{\frac{-2\rho}{2+\rho}} + 0.5 \right)^{\frac{1+\rho}{-\rho}}$, $e_2^N(\rho, \epsilon) =$

$\dfrac{1}{\eta} \left(\dfrac{3-\epsilon}{6} \right)^2 \left(0.5 \left(\dfrac{3+\epsilon}{3-\epsilon} \right)^{\frac{-2\rho}{2+\rho}} + 0.5 \right)^{\frac{1+\rho}{-\rho}}$, and $\delta^N(\rho, \epsilon) = \dfrac{1}{36\eta} \times$

$\left(0.5 \, (3+\epsilon)^{\frac{-2\rho}{2+\rho}} + 0.5 \, (3-\epsilon)^{\frac{-2\rho}{2+\rho}} \right)^{\frac{2+\rho}{-\rho}}$. These expressions include the limit cases

of $e_i^N(-1, \epsilon) = \dfrac{1}{\eta} \left(\dfrac{3+\epsilon_i}{6} \right)^2$ $(i = 1, 2, \ \epsilon_1 = \epsilon = -\epsilon_2)$, $\delta^N(-1, \epsilon) = \dfrac{9+\epsilon^2}{36\eta}$,

$e_i^N(0, \epsilon) = \dfrac{(3+\epsilon_i)^{3/2} (3-\epsilon_i)^{1/2}}{36\eta}$, $\delta^N(0, \epsilon) = \dfrac{(3+\epsilon)(3-\epsilon)}{36\eta}$ and $e_1^N(\infty, \epsilon) =$

$\dfrac{(3-\epsilon)^2}{36\eta} \dfrac{(3+\epsilon)^2}{\epsilon^2 + 9} = e_2^N(\infty, \epsilon) = \delta^N(\infty, \epsilon)$. From these equations, we obtain $e_1^N(\rho, \epsilon) > e_2^N(\rho, \epsilon) > 0$ for all $\rho < \infty$ and $\epsilon > 0$.

Next, we show that $e_2^N(\rho, \epsilon)$ is increasing in ρ. We begin with the case of $\rho \neq 0$. Setting $x \equiv \left(\dfrac{3+\epsilon}{3-\epsilon} \right)^2 \in [1, 4)$ and $y \equiv \dfrac{\rho}{2+\rho} \in [-1, 1) \setminus \{0\}$, we have:

$$\dfrac{\partial e_2^N(\rho, \epsilon)}{\partial \rho} = \dfrac{1}{\eta} \left(\dfrac{3-\epsilon}{6} \right)^2 \dfrac{(-1+y)^2 (0.5x^{-y} + 0.5)^{-\frac{1+3y}{2y}}}{4y^2} \times$$
$$\left((0.5x^{-y} + 0.5) \ln (0.5x^{-y} + 0.5) - (0.5 + 0.5y) x^{-y} \ln \left(x^{-y} \right) \right).$$

Let $Z(x, y) \equiv \left((0.5x^{-y} + 0.5) \ln (0.5x^{-y} + 0.5) - (0.5 + 0.5y) x^{-y} \ln (x^{-y}) \right)$. We have $Z(1, y) = 0$ for all y. For $x > 1$ and $z \equiv x^{-y} \neq 1$, $Z(x, -\ln(z)/\ln(x)) > 0$ if and only if $\ln x < \dfrac{(\ln z)^2}{\ln z - (z^{-1} + 1) \ln(0.5z + 0.5)} \equiv f(z)$. Indeed, $f(z) > \ln 4$ for all $z \in (0.25, 4) \setminus \{1\}$. Therefore, in the range of $x \in (1, 4)$ and $y \in [-1, 1) \setminus \{0\}$, we have $Z(x, y) > 0$, so that $\partial e_2^N(\rho, \epsilon)/\partial \rho > 0$ for all $\rho \geq -1$ and $\epsilon > 0$.

When $\rho = 0$, $\dfrac{\partial e_2^N(\rho, \epsilon)}{\partial \rho} = \dfrac{1}{36\eta} (3-\epsilon)^2 \dfrac{1}{8} x^{\frac{1}{4}} \ln x \left(1 - \dfrac{1}{4} \ln x \right) > 0$ for $x \in (1, 4)$. Likewise, one obtains $\partial e_1^N(\rho, \epsilon)/\partial \rho < 0$ for all $\rho \geq -1$ and $\epsilon > 0$.

Next, we show that $e_2^N(\rho, \epsilon)$ is decreasing in ϵ. As $e_2^N(\rho, \epsilon) = (36\eta)^{\frac{-1}{2+\rho}} \times (\delta^N(\rho, \epsilon))^{\frac{1+\rho}{2+\rho}} (3-\epsilon)^{\frac{2}{2+\rho}}$ and we show below that $\delta^N(\rho, \epsilon)$ is decreasing in ϵ when $\rho > -\dfrac{2}{3}$, it is sufficient to see if $e_2^N(\rho, \epsilon)$ is decreasing in ϵ when $\rho \leq -\dfrac{2}{3}$. Setting x as before and $y = \dfrac{\rho}{2+\rho} \in [-1, -0.5]$, we have:

$$\mathrm{sgn} \left(\dfrac{\partial e_2^N(\rho, \epsilon)}{\partial \epsilon} \right) = \mathrm{sgn} \left(x^{-y} (x^{\frac{1}{2}} y - x^{\frac{1}{2}} + y + 1) - 2x^{\frac{1}{2}} \right).$$

22

$x^{\frac{1}{2}}y - x^{\frac{1}{2}} + y + 1 < 0$ if $y < \dfrac{x^{\frac{1}{2}} - 1}{x^{\frac{1}{2}} + 1}$, which is the case for $x \in [1, 4)$ and $y \in [-1, -0.5]$. Therefore, $e_2^N(\rho, \epsilon)$ is decreasing in ϵ for all ϵ. The sign of $\partial e_1^N(\rho, \epsilon)/\partial \epsilon$ is ambiguous, as Figure 3 indicates that $e_1^N(\rho, \epsilon)$ may be higher or lower with a greater ϵ.

Differentiating $e_1^N(\rho, \epsilon) - e_2^N(\rho, \epsilon)$ with respect to ϵ, we obtain:

$$\mathrm{sgn}\left(\frac{\partial(e_1^N(\rho, \epsilon) - e_2^N(\rho, \epsilon))}{\partial \epsilon}\right) = \mathrm{sgn}\left\{\frac{(3-\epsilon)^{\frac{2}{2+\rho}}}{(3+\epsilon)^{\frac{3\rho+2}{2+\rho}}}\left((2+\rho)\left(\frac{3+\epsilon}{3-\epsilon}\right)^{\frac{2}{2+\rho}} - \rho + \frac{2\epsilon}{3-\epsilon}\right)\right.$$
$$\left. + \frac{(3+\epsilon)^{\frac{2}{2+\rho}}}{(3-\epsilon)^{\frac{3\rho+2}{2+\rho}}}\left((2+\rho)\left(\frac{3-\epsilon}{3+\epsilon}\right)^{\frac{2}{2+\rho}} - \rho - \frac{2\epsilon}{3+\epsilon}\right)\right\}.$$

We can show numerically that this is positive. Evaluated at $\rho = 0$, we also have $\partial(e_1^N(\rho, \epsilon) - e_2^N(\rho, \epsilon))/\partial\epsilon = (18 - 4\epsilon^2)/(36\eta\sqrt{3-\epsilon}\sqrt{3+\epsilon}) > 0$.

Differentiating $\delta^N(\rho, \epsilon)$ with respect to ρ, we obtain:

$$\frac{\partial \delta^N(\rho, \epsilon)}{\partial \rho} = \frac{1}{36\eta}\left(0.5\,(3+\epsilon)^{\frac{-2\rho}{2+\rho}} + 0.5\,(3-\epsilon)^{\frac{-2\rho}{2+\rho}}\right)^{\frac{2+\rho}{-\rho}-1}\frac{2}{\rho^2} \times$$
$$\left\{\left(0.5\,(3+\epsilon)^{\frac{-2\rho}{2+\rho}} + 0.5\,(3-\epsilon)^{\frac{-2\rho}{2+\rho}}\right)\ln\left(0.5\,(3+\epsilon)^{\frac{-2\rho}{2+\rho}} + 0.5\,(3-\epsilon)^{\frac{-2\rho}{2+\rho}}\right)\right.$$
$$\left. -0.5\,(3+\epsilon)^{\frac{-2\rho}{2+\rho}}\ln(3+\epsilon)^{\frac{-2\rho}{2+\rho}} - 0.5\,(3-\epsilon)^{\frac{-2\rho}{2+\rho}}\ln(3-\epsilon)^{\frac{-2\rho}{2+\rho}}\right\} < 0,$$

from Jensen's inequality because the function $g(a) = a\ln a$ is convex with respect to a. Evaluated at $\rho = 0$, we also have $\partial\delta^N(\rho, \epsilon)/\partial\rho = -(3 - \epsilon)(3 + \epsilon)(\ln(3 + \epsilon) - \ln(3 - \epsilon))^2/(144\eta) < 0$.

Differentiating $\delta^N(\rho, \epsilon)$ with respect to ϵ at $\epsilon > 0$, we obtain:

$$\frac{\partial \delta^N(\rho, \epsilon)}{\partial \epsilon} = \frac{1}{36\eta}\left(0.5\,(\epsilon+3)^{\frac{-2\rho}{2+\rho}} + 0.5\,(3-\epsilon)^{\frac{-2\rho}{2+\rho}}\right)^{\frac{2+\rho}{-\rho}-1}\left((3+\epsilon)^{\frac{-3\rho-2}{2+\rho}} - (3-\epsilon)^{\frac{-3\rho-2}{2+\rho}}\right),$$

which is positive if $\rho < -\dfrac{2}{3}$ and negative if $\rho > -\dfrac{2}{3}$. For $\rho = 0$, $\dfrac{\partial \delta^N(\rho, \epsilon)}{\partial \epsilon} = -\dfrac{\epsilon}{18\eta}$. For $\rho = -\dfrac{2}{3}$, $\delta^N\left(-\dfrac{2}{3}, \epsilon\right) = \dfrac{1}{4\eta}$ for all $\epsilon \geq 0$. *Q.E.D.*

Proof of Proposition 1: The first-order conditions of the joint welfare maximization (9) are $e_i^{-\rho-1}\left(0.5e_1^{-\rho} + 0.5e_2^{-\rho}\right)^{\frac{1+\rho}{-\rho}}\left(\dfrac{9+\epsilon^2}{18}\right) - \eta e_i = 0$ ($i = 1, 2$), from which we obtain $e_1^{N*}(\rho, \epsilon) = e_2^{N*}(\rho, \epsilon) = \dfrac{9+\epsilon^2}{18\eta}$. When $\rho = -1$, $W_2^N(e^{N*}(-1, \epsilon)) - W_2^N(e^N(-1, \epsilon)) = \dfrac{81 + \epsilon^4 - 36\epsilon^3 + 54\epsilon^2 - 324\epsilon}{2 \cdot 36^2\eta}$, which is

23

positive when $\epsilon < 9+6\sqrt{2}-6\sqrt{4+3\sqrt{2}} \equiv \epsilon^N(-1) \approx 0.2592817$ and negative when $\epsilon > \epsilon^N(-1)$. For $\rho > -1$, $W_2^N(e^{N*}(\rho,\epsilon)) = \dfrac{\delta(e^{N*}(\rho,\epsilon))}{2}\left(\dfrac{3-\epsilon}{3}\right)^2 - \eta\dfrac{(e_2^{N*}(\rho,\epsilon))^2}{2}$ is invariant with respect to ρ, whereas $W_2^N(e^N(\rho,\epsilon)) = \dfrac{\delta^N(\rho,\epsilon)}{2}\left(\dfrac{3-\epsilon}{3}\right)^2 - \eta\dfrac{(e_2^N(\rho,\epsilon))^2}{2}$ decreases in ρ from Lemma 4. We numerically show that there exists an increasing function $\epsilon^N(\rho)$ such that $W_2^N(e^{N*}(\rho,\epsilon)) - W_2^N(e^N(\rho,\epsilon))$ is positive (negative) when $\epsilon < \epsilon^N(\rho)$ ($\epsilon > \epsilon^N(\rho)$). $\epsilon^N(\rho)$ is shown in Figure 4: for example, $\epsilon^N(-0.2) \approx 0.3114695738$, $\epsilon^N(0) \approx 0.3195211862$, and $\epsilon^N(\infty) \approx 0.4094092130$.

For country 1, from Lemma 4, we have $W_1^N(e^N(\rho,\epsilon)) < \dfrac{\delta^N(-1,\epsilon)}{2}\left(\dfrac{3+\epsilon}{3}\right)^2 - \eta\dfrac{(e_1^N(\infty,\epsilon))^2}{2} \equiv w(\epsilon)$ for all ρ. $W_1^N(e^{N*}(\rho,\epsilon)) - w(\epsilon) = $
$$\dfrac{\epsilon^8 + 24\epsilon^7 + 36\epsilon^5(1-\epsilon) + 612\epsilon^5 + 486\epsilon^4 + 5832\epsilon^3 + 2916\epsilon(1-\epsilon) + 14580\epsilon + 6561}{2 \cdot 36^2\eta(\epsilon^2+9)^2} >$$
0 for all ρ. Therefore, we have $W_1^N(e^N(\rho,\epsilon)) < w(\epsilon) < W_1^N(e^{N*}(\rho,\epsilon))$ for all ρ and ϵ. $Q.E.D.$

Proof of Lemma 5: From (5), the tax reaction function of country 2 is $\arg\max_{t_2} R_2(t_2,t_1,e) \equiv \hat{t}_2(t_1;e) = \delta(e)\left(\dfrac{1-\epsilon}{2}\right) + \dfrac{t_1}{2}$. We have $\dfrac{\partial\hat{t}_2}{\partial t_1} \in (0,1)$. The first-order condition of the tax leader is given by $\dfrac{\partial W_1}{\partial t_1} + \dfrac{\partial W_1}{\partial t_2}\dfrac{\partial\hat{t}_2}{\partial t_1} = \dfrac{1+\epsilon}{2} - \dfrac{t_1-\hat{t}_2(t_1)}{2\delta(e)} - t_1\dfrac{1-(1/2)}{2\delta(e)} = 0,$[25] which yields $t_1^S(e) = \delta(e)\left(\dfrac{3+\epsilon}{2}\right)$ and $t_2^S(e) = \delta(e)\left(\dfrac{5-\epsilon}{4}\right)$. From (6), we have $t_i^S(e) > t_i^N(e)$ $(i = 1,2)$ and $t_1^S(e) - t_2^S(e) = \dfrac{\delta(e)}{4}(1+3\epsilon) > \dfrac{2\delta(e)\epsilon}{3} = t_1^N(e) - t_2^N(e)$, as in parts (i) and (ii) of the lemma. For $\tilde{\pi}_i^S \equiv \tilde{\pi}_i(t_1^S(e),t_2^S(e),e)$, from Lemma 1, we have $\tilde{\pi}_1^S = \dfrac{1+\epsilon}{2} - \dfrac{t_1^S(e)-t_2^S(e)}{2\delta(e)} = \dfrac{3+\epsilon}{8} < \dfrac{5-\epsilon}{8} = \tilde{\pi}_2^S$. Then, we have $R_1^S(e) = t_1^S(e)\tilde{\pi}_1^S = \dfrac{\delta(e)}{16}(3+\epsilon)^2$ and $R_2^S(e) = \dfrac{\delta(e)}{32}(5-\epsilon)^2$. From (7), $R_1^S(e) - R_1^N(e) = \dfrac{\delta(e)}{16}(3+\epsilon)^2 - \dfrac{\delta(e)}{18}(3+\epsilon)^2 > 0$ and $R_2^S(e) - R_2^N(e) = $

---

[25]Note that the revenue function is concave in the tax rate, $\partial\left(\dfrac{\partial W_1}{\partial t_1} + \dfrac{\partial W_1}{\partial t_2}\dfrac{\partial\hat{t}_2}{\partial t_1}\right)/\partial t_1 = -1/(2\delta(e)) < 0$.

$\frac{\delta(e)}{32}(5-\epsilon)^2 - \frac{\delta(e)}{18}(3-\epsilon)^2 = \frac{\delta(e)}{288}(3+\epsilon)(27-7\epsilon) > 0$ in part (ii). In addition, $R_1^S(e) - R_2^S(e) = \frac{\delta(e)}{32}(-7 + \epsilon^2 + 22\epsilon) \gtreqless 0 \iff \epsilon \gtreqless 8\sqrt{2} - 11 \equiv \hat{\epsilon} \approx 0.31370850$. For $\epsilon < \hat{\epsilon}$, $0 < R_2^S(e) - R_1^S(e) \lesseqgtr R_1^N(e) - R_1^N(e) \iff \epsilon \gtreqless \tilde{\epsilon} \approx 0.16094072$. For all $\epsilon \geq \hat{\epsilon}$, we have $0 \leq R_1^S(e) - R_2^S(e) < R_1^N(e) - R_2^N(e)$. Therefore, part (iii) of the lemma holds. *Q.E.D.*

Proof of Lemma 6: From Lemma 5 and (10), $W_1^S(e_1, e_2) = \frac{0.5e_1 + 0.5e_2}{16}(3 + \epsilon)^2 - \eta\frac{(e_1)^2}{2}$. The first-order condition with respect to e_1 yields $e_1^S = \frac{(3+\epsilon)^2}{32\eta}$. Likewise, $e_2^S = \frac{(5-\epsilon)^2}{64\eta}$. From Lemma 3, $e_1^N = e_1^N(-1, \epsilon) = \frac{(3+\epsilon)^2}{36\eta}$ and $e_2^N = \frac{(3-\epsilon)^2}{36\eta}$, so we have $e_1^S > e_1^N$, $e_2^S > e_2^N$ and $\delta^S > \delta^N$. As in Lemma 5.(iii), $|e_2^S - e_1^S| < e_1^N - e_2^N$ if and only if $\epsilon > \tilde{\epsilon}$. Therefore, parts (i) and (iii) of the lemma hold. $\delta^S > \delta^N$ and Lemma 5 imply that $R_i^S(e^S) = t_i^S(e^S)\tilde{\pi}_i^S > t_i^S(e^N)\tilde{\pi}_i^S = R_i^S(e^N) > R_i^N(e^N)$ $(i = 1, 2)$ and $t_1^S(e^S) - t_2^S(e^S) > t_1^S(e^N) - t_2^S(e^N) > t_1^N(e^N) - t_2^N(e^N)$, so part (ii) of the lemma holds. *Q.E.D.*

Proof of Proposition 2: From Lemma 5, the first-order conditions of the joint welfare maximization $\max_{e_1, e_2} \sum_i \left(R_i^S(e_i, e_j) - \eta\frac{(e_i)^2}{2} \right)$ are $\frac{(3+\epsilon)^2}{32} + \frac{(5-\epsilon)^2}{64} - \eta e_i = 0$ $(i = 1, 2)$, from which we obtain $e_1^{S*} = e_2^{S*} = \frac{3\epsilon^2 + 2\epsilon + 43}{64\eta}$. Let $W_i^S(e^{S*})$ and $W_i^S(e^S)$ $(i = 1, 2)$ be the welfare levels in the noncooperative and cooperative regimes, respectively. For the tax leader, $W_1^S(e^{S*}) - W_1^S(e^S) = \frac{23 + 7\epsilon^4 + 116\epsilon^3 + 282\epsilon^2 + 1364\epsilon}{2 \cdot 64^2 \eta} > 0$ for all ϵ. For the tax follower, $W_2^S(e^{S*}) - W_2^S(e^S) = \frac{463 - \epsilon^4 - 44\epsilon^3 + 42\epsilon^2 - 716\epsilon}{64^2 \eta}$, which is positive when $\epsilon < 8\sqrt{4 + 3\sqrt{2}} - 11 - 8\sqrt{2} \equiv \epsilon^S \approx 0.6542911$ and negative when $\epsilon > \epsilon^S$. We have $\epsilon^S > \epsilon^N(-1)$. *Q.E.D.*

References

[1] Altshuler, R. and Goodspeed, T. 2015. Follow the leader? Evidence on European and US tax competition, *Public Finance Review* 43, 485–504.

[2] Amerighi, O., Peralta, S. 2010. The proximity-concentration trade-off with profit shifting, *Journal of Urban Economics* 68, 90–101.

[3] Bacchetta, P., Espinosa, M. P., 1995. Information sharing and tax competition among governments, *Journal of International Economics* 39, 103–121.

[4] Bartelsman, E., Beetsma, R., 2003. Why pay more? Corporate tax avoidance through transfer pricing in OECD countries, *Journal of Public Economics* 87, 2225–2252.

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

[6] Bucovetsky, S., Haufler, A., 2008. Tax competition when firms choose their organizational form: Should tax loopholes for multinationals be closed? *Journal of International Economics* 74, 188–201.

[7] Chatelais, N., Peyrat, M., 2008. Are small countries leaders of the European tax competition? Working paper 2008.58, Centre d'Economie de la Sorbonne, Paris.

[8] Clausing, K., 2003. Tax motivated transfer pricing and US intrafirm trade prices, *Journal of Public Economics* 87, 2207–2223.

[9] Cremer, H., Gahvari, F., 2000. Tax evasion, fiscal competition and economic integration, *European Economic Review* 44, 1633–1657.

[10] Devereux, M., Lockwood, B., Redoano, M., 2008. Do countries compete over corporate tax rates? *Journal of Public Economics* 92, 1210–1235.

[11] Durán-Cabré, J., Esteller-Moré, A., Salvadori, L., 2015. Empirical evidence on horizontal competition in tax enforcement, *International Tax and Public Finance* 22, 834–860.

[12] European Commission, 2018. *The Fiscalis 2020 programme,* Taxation and Customs Union, European Commission, Brussels, https://ec.europa.eu/taxation_customs/fiscalis-programme_en (accessed 25 September 2018).

[13] Gordon, R., 1992. Can capital income taxes survive in open economies? *Journal of Finance* 47, 1159–1180.

[14] Grubert, H., Altshuler, R., 2013. Fixing the system: An analysis of alternative proposals for the reform of international tax, *National Tax Journal* 66, 671–712.

[15] Haufler, A., Schjelderup, G., 2000. Corporate tax systems and cross country profit shifting, *Oxford Economic Papers* 52, 306–325.

[16] Hamilton, J., Slutsky, S., 1990. Endogenous timing in duopoly games: Stackelberg or Cournot equilibria, *Games and Economic Behavior* 2, 29–46.

[17] Hindriks, J., Nishimura, Y., 2015. A note on equilibrium leadership in tax competition models, *Journal of Public Economics* 121, 66–68.

[18] Hindriks, J., Nishimura, and Y., 2017. Equilibrium leadership in tax competition models with capital ownership: A rejoinder, *International Tax and Public Finance* 24, 338–349.

[19] Hindriks, J., Peralta, S., Weber, S., 2014. Local taxation of global corporation: A simple solution, *Annals of Economics and Statistics* 113–114, 37–65.

[20] Hirshleifer, J., 1983. From weakest-link to best-shot: The voluntary provision of public goods, *Public Choice* 41, 371–386.

[21] Huizinga, H., Laeven, L., 2008. International profit shifting within multinationals: A multi-country perspective, *Journal of Public Economics* 92, 1164–1182.

[22] Huizinga, H., Laeven, L., Nicodeme, G., 2008. Capital structure and international debt shifting, *Journal of Financial Economics* 88, 80–118.

[23] Kanbur, R., Keen, M., 1993. Jeux sans frontieres: Tax competition and tax coordination when countries differ in size, *American Economic Review* 83, 877–892.

[24] Keen, M., Konrad, K., 2013. The theory of international tax competition and coordination, in Auerbach, A., Chetty, R., Feldstein, M., Saez, E., (Eds.), *Handbook of Public Economics*, volume 5, North Holland, Amsterdam, pp. 257–328.

[25] Kempf, H., Rota-Graziosi, G., 2010. Endogenizing leadership in tax competition, *Journal of Public Economics* 94, 768–776.

[26] Kind, H.J., Midelfart, K. H., Schjelderup, G., 2005. Corporate tax systems, multinational enterprises and economic integration, *Journal of International Economics* 65, 507–521.

[27] Klassen, K., Laplante, S., 2012. Are U.S. multinational corporations becoming more aggressive income shifters?, *Journal of Accounting Research* 50, 1245–1285.

[28] Konrad, K., 2008. Mobile tax base as a global common, *International Tax and Public Finance* 15, 395–414.

[29] Lohse, T., Riedel, N., 2013. Do transfer pricing laws limit international income shifting? Evidence from European multinationals, *CESifo Working Paper* 4404, CESifo GmbH, Munich.

[30] Mintz, J., Smart, M., 2004. "Income shifting, investment and tax competition: Theory and evidence from provincial taxation in Canada", *Journal of Public Economics* 88, 1149–1168.

[31] Næss-Schmidt, P., Sunesen, E., Hansen, M., Ali, D., 2012. Tax elasticities of financial instruments, profits and remuneration, Working Paper N.30-2012, European Commission, Luxembourg.

[32] Nielsen, S. B., Raimondos-Moller, P., Schjelderup, G., 2008. Taxes and decision rights in multinationals, *Journal of Public Economic Theory* 10, 245–258.

[33] OECD, 2015. *Base erosion and profit shifting,* Organisation for Economic Co-operation and Development, Paris, http://www.oecd.org/tax/beps.htm (accessed 25 September 2018).

[34] Peralta, S., Wauthy, X., van Ypersele, T., 2006. Should countries control international profit shifting?, *Journal of International Economics* 68, 24–37.

[35] Stöwhase, S., 2013. How profit shifting may increase the tax burden of multinationals: A simple model with discrete investment choices, *Journal of Public Economic Theory* 15, 185–207.

[36] Stöwhase, S., Traxler., C., 2005. Tax evasion and auditing in a federal economy, *International Tax and Public Finance* 12, 515–531.

[37] Swenson, D., 2001. Tax reforms and evidence of transfer pricing, *National Tax Journal* 54, 7–25.

[38] Wang, Y.-Q., 1999. Commodity taxes under fiscal competition: Stackelberg equilibrium and optimality, *American Economic Review* 89, 974–981.

[39] Wilson, J., 1991. Tax competition with interregional differences in factor endowments, *Regional Science and Urban Economics* 21, 423–451.