# Discussion Papers In Economics And Business

Enforcement Cooperation and Tax Competition Among Asymmetric Countries

Jean HINDRIKS and Yukihiro NISHIMURA

Discussion Paper 17-23

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

# Enforcement Cooperation and Tax Competition Among Asymmetric Countries

Jean HINDRIKS and Yukihiro NISHIMURA

Discussion Paper 17-23

July 2017

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

# Enforcement Cooperation and Tax Competition Among Asymmetric Countries*

Jean HINDRIKS[1] and Yukihiro NISHIMURA[2]

July 26, 2017

**Abstract**

Using a profit shifting model with multinational enterprises that operate in two countries, large and small, we analyze the determinants of voluntary cooperation on enforcement effort and information sharing arrangements. The benefit from tougher enforcement by one country is the mitigation of the tax competition, so that enforcement choice has a public-good nature. In a framework that countries decide tax rates noncooperatively, we compare the equilibria of the noncooperative and cooperative choice of enforcement. With sufficient disparity in the country sizes, the low-tax country is not willing to participate enforcement coordination. We then consider two extensions. The first extension is complementarity (imperfect substitutability) of countries' enforcement efforts, which comes from administrative or legal nature of the collective decision-making. We show that cooperation is more viable with greater enforcement complementarity. In the second extension, we consider the Stackelberg tax competition where the large country leads. Compared with the simultaneous tax choice, both the tax leader and the tax follower will exert more enforcement efforts in the sequential tax choice. As a result, the enforcement cooperation will be more viable under the Stackelberg tax competition.

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

**Keywords:** Profit shifting, Tax competition, Information sharing, Tax leadership

---

*Yukihiro Nishimura acknowledges the financial support from the Grants-in-Aid for Scientific Research (C) from the Ministry of Education, Culture, Sports, Science, and Technology (24530348, 15K03511) and Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation from the Japan Society for the Promotion of Science (J2402).

[1]CORE, Université catholique de Louvain. jean.hindriks@uclouvain.be

[2]Graduate School of Economics, Osaka University. 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 source-based taxation (also known as territorial system),[1] these MNEs found that they can reduce their overall tax liabilities by shifting profit and income between branches, which caused a well-known problem of Base Erosion and Profit Shifting (BEPS, hereafter referred to as "profit shifting"). The presence of profit shifting gets convincing empirical support.[2] To address this issue, OECD's (2015) report proposed global action plans, which contains a series of specific enforcement efforts to be taken by the governments.[3] However, a real challenge to international coordination is the absence of enforceable global institution, and, in addition, the difference across countries regarding preferred tax rates on the corporate incomes. Indeed, in reality, the enforcement of legislations differ significantly across countries: some legislations 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 on enforcement effort and information sharing arrangements. To address the present issues in a tax competition framework, we develop a simple two-country model with different market sizes, which follows Kanbur and Keen (1993), Hindriks et al. (2014) and Keen and Konrad (2013). MNEs shift

[1] Most OECD countries adopted territorial systems, since it is relatively easy for taxpayers to avoid resident-based taxes on capital income. Bacchetta and Espinosa (1995) showed that countries do not choose the residence-based principle when information transmission can be strategic. Profit shifting is also observed under the tax system of the United States; see, for example, Swenson (2001), Clausing (2003), Klassen and Laplante (2012), and Grubert and Altshuler (2013).

[2] Swenson (2001) and Clausing (2003) found that taxation has a significant effect on intra-firm prices of the MNEs in the United States. Mintz and Smart (2003) found that Canadian firms which operate in multiple jurisdictions have a high elasticity of taxable income with respect to corporate tax rates. Bartelsmann and Beetsma (2003) used the 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. Huizinga and Laeven (2008) examined European MNEs and found that profit shifting leads to a substantial redistribution of corporate tax revenues. IMF (2011) used Bankscope data on banks to find that 1 percent-point higher tax rate reduces reported fiscal profits by between 6 and 8.5 percent. In the financial sector, Næss-Schmidt et al. (2012) found that a one percent increase in the tax rate leads to a 1.9 percent decrease in an MNE's affiliates. See also Devereux and Griffith (2003) and Auerbach et al. (2010).

[3] These proposals include countering harmful tax practices (Action 5), mandatory disclosure of suspected aggressive tax planning (Action 12), transfer pricing documentation and country-by-country reporting (Action 13), and effectiveness of dispute resolution (Action 14).

profits from the high to the low tax country's division, subject to a concealment cost. Countries choose the level of enforcement efforts, such as more efficient information sharing, strict monitoring and inspection, and the effort for negotiation and agreement with the other country's tax authority. The enforcement efforts incur administration costs to each country. But its benefit, the reduction of aggressive tax planning that mitigates the tax competition, is diffused to the other country. Therefore, enforcement choice becomes a game of voluntary contribution to a public good. We compare the equilibria of the noncooperative and cooperative choice of enforcement. In the latter scenario, countries choose the enforcement levels so as to maximize their joint welfare, but countries still decide tax rates noncooperatively.[4] In the absence of lump-sum transfers to share the gain of cooperation, with sufficient disparity in the country sizes, the low-tax country is not willing to participate enforcement coordination (Proposition 1). It is of central concern in the application of coordination proposals to OECD, G20 and the European Union (EU), since the adoption of such proposals often requires unanimity.

We will then identify two different drivers of cooperation. The first driver is complementarity in the enforcement choices. Enforcement efforts increase the cost of profit shifting. However, in reality, dispersed (unilateral) enforcement efforts by the involved countries are not effective. For instance, the lack of exchange of tax-relevant information by host country makes the taxable income unclear to the home country. Indeed, for recommendations of the OECD BEPS Project, including mandatory disclosure of cross-border tax planning (Action 12) and advance pricing arrangement (Action 14), countries are free to choose whether to introduce them or not, based on each country's domestic tax regime.[5] Also, during the mutual agreement procedure on transfer prices and taxable incomes of the MNEs, the capital-importing country may well be inclined to favor the MNE's financial and legal expert's view towards lower transfer prices and higher taxable incomes to its own. Given that double taxation is not allowed by the tax treaty, the capital-importing country can thus exercise a veto power towards tighter enforcement.[6] To formalize the notion of imperfect substitutability (complementarity), we apply Hirshleifer's (1983) social composition function. The case of veto power corresponds to the weakest-link formula, where the capital-exporting country has to compromise to the minimum of the enforcement standard. If in contrast,

[4]Host country retains the primary taxing rights over source-based incomes. Accordingly, preferred tax rates differ across countries.

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

[6]This scenario in a collective decision-making was mentioned by Hirshleifer (1983, p. 373).

relevant information including the transfer pricing documentation from each country (Article 13 of OECD (2015)) can be added up towards better tax enforcement, then the associated efforts have more substitutable (additive) nature. By parametrically treating the degree of complementarity, we show that cooperation is more viable with greater enforcement complementarity (Proposition 2).

The second driver of cooperation is tax leadership. The analysis of endogenous leadership has attracted research interest since Kempf and Rota-Graziosi (2010).[7] Following Hindriks and Nishimura (2015), we consider the case in which the large country leads.[8] Compared with simultaneous tax choice, sequential tax choice makes both the tax leader and the tax follower exert more enforcement effort. Moreover, the Stackelberg tax competition reduces the dispersion of the enforcement efforts, provided that asymmetry is sufficiently large (Lemma 5). As a result, the enforcement cooperation will be more viable than under the Nash tax competition (Proposition 3).

The main interest of the previous studies was to understand how transfer prices are affected by international tax differences and tax systems (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)). Some evidence also showed that transfer pricing regulations significantly mitigate profit shifting (for example, Bartelsmann and Beetsma (2003) and Lohse and Riedel (2013)). However, these studies did not examine the choice (determinants) of enforcement efforts by countries with different economic structure.[9] In a recent paper, Hindriks et al. (2014) proposed a solution based on a system of voluntary tax sharing agreement among asymmetric countries. In this paper, we endogenize the MNE's concealment cost and examine policy issues with respect to enforcement complementarity and tax leadership. Cremer and Gahvari (2000) and Stöwhase and Traxler (2005) examined the implications of tax evasion for

[7] Altshuler and Goodspeed (2015) demonstrated that European countries set their corporate tax rates following the United States 1986 Tax Reform Act. Also, Redoano (2007) and Chatelais and Peyrat (2008) provided evidence of sequential tax decisions among European countries. Stöwhase (2013) adopted Stackelberg setting, simply to avoid the difficulties resulting from possible discontinuities in the payoff functions with simultaneous tax setting. As such, in his framework, the Nash and the Stackelberg tax competitions are not comparable. Also, unlike ours, a public-good nature of the tax enforcement is absent.

[8] In our model, the large leadership equilibrium Pareto dominates the small leadership equilibrium under sufficient asymmetry, as in Hindriks and Nishimura (2015).

[9] Exceptions are Bacchetta and Espinosa (1995), Peralta et al. (2006), and Bucovetsky and Haufler (2008). However, our formal model differs from these papers. More importantly, none of these papers discussed the determinants of the emergence of voluntary cooperation on enforcement effort.

tax competition. However, unlike ours, the tax evasion and the tax enforcement (auditing) in their papers do not have horizontal interactions.[10] In contrast, in the BEPS problem, the shifted incomes from one branch (net of the concealment cost) become 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 between regional administrations in 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. Section 4 examines two extensions, enforcement complementarity and tax leadership. Section 5 concludes. The proofs of several 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 is characterized by the linear (inverse) demands 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

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

Namely, Country 1 is the *large* country, either because (i) $\alpha_i$ that represents population size or income per capita is higher, or because (ii) the supply cost $c_i$ is lower.

Associated with 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 it may, at some cost, shift profits between branches so as to minimize their total tax liability. Namely, 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, the 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 follow Haufler and

[10] In Cremer and Gahvari (2000) and Stöwhase and Traxler (2005), a firm decides on concealment depending solely on the tax rates of the home country, and the convex auditing cost does not depend on the intensity of auditing of the other countries.

Schjelderup's (2000) transfer-pricing model and introduce a convex and non-fiscally deductible concealment cost $C\left(\pi_i^k, \tilde{\pi}_i^k\right)$ (widely used by, for example, Swenson (2001), Kind et al. (2005), Peralta et al. (2006), Devereux et al. (2008), Nielsen et al. (2008), and Keen and Konrad (2013)):[11]

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

Several explanations are in order. First, $\delta(e)$ is a scaling factor of resource costs associated with profit shifting or transfer pricing. It reflects the cost of hiring accounting experts, or the expected market sanction when caught cheating on tax liabilities. As a standard assumption in the literature, the concealment cost is an increasing and convex function of the amount of profit shifting, $|\pi_i^k - \tilde{\pi}_i^k|$, independent of the direction of shifting (outward and inward shiftings are cost-equivalent).

Second, $\delta(e) = \delta(e_i, e_j)$ depends on the governments' enforcement efforts $e_i, e_j$, such as more efficient information sharing, tougher monitoring, and the effort for negotiation and agreement with the other country's tax authority. $\delta(e)$ is an increasing function of $e_i$ and $e_j$, so that stricter enforcement implies higher $\delta(e)$. It can either be additively separable or there is interdependence between $e_i$ and $e_j$. In the former case, the cost (1) can be interpreted as the tax compliance costs separately required by the two governments (for example, documentation requirements).[12]

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 the consumer surplus is independent of the tax choices. This feature is synonymous to widely used model by Kanbur and Keen (1993). We assume

[11]See also Huizinga and Laeven (2008) and Amerighi and Peralta (2010) for a slightly different specification. We do not expect our results to be influenced by this alternative specification. In a related model of international debt shifting by MNEs, Huizinga et al. (2008) considered the model of allocating tax allowances of debt finance, whereby allocation of the debt affects the expected bankruptcy costs which are not tax-deductible. As in Lemma 1 below, the firm utilizes international tax rate differences on the debt allocation. The concealment cost may reflect penalties to be paid to the government. However, as a standard assumption in this literature, we consider the scenario that the concealment is a pure waste of resources.

[12]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.

that $t_i \leq 1$, for $i = 1, 2$. Assuming quadratic cost of enforcement ($c(e_i) = \eta\frac{e_i^2}{2}$), welfare in country $i$ is

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

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

## 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 profit to be shifted.

The issue of tax and enforcement timing will be described and discussed in the next subsection. In this subsection, we first analyze firms' decisions 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, $(\tilde{\pi}_1^k, \tilde{\pi}_2^k)$, so as to maximize the after-tax profit net of the shifting cost

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

subject to $\tilde{\pi}_1^k + \tilde{\pi}_2^k = \pi_1^k + \pi_2^k$. Let $\gamma_1 = \frac{3}{2}\sqrt{\beta(1+\epsilon)}$, $\gamma_2 = \frac{3}{2}\sqrt{\beta(1-\epsilon)}$ with $\epsilon \in [0, 1)$, and let $\tilde{\pi}_i = \tilde{\pi}_i^a + \tilde{\pi}_i^b$ be the total reported profit in country $i$. Hindriks et al. (2014) showed the following:[13]

**Lemma 1** $\pi_1^k = \dfrac{\gamma_1^2}{9\beta} \equiv \dfrac{1+\epsilon}{4}$, $\pi_2^k = \dfrac{1-\epsilon}{4}$ $(k = a,\ b)$, $\tilde{\pi}_1 = \tilde{\pi}_1^a + \tilde{\pi}_1^b = \dfrac{1+\epsilon}{2} - \dfrac{t_1 - t_2}{2\delta(e)}$, $\tilde{\pi}_2 = \tilde{\pi}_2^a + \tilde{\pi}_2^b = \dfrac{1-\epsilon}{2} + \dfrac{t_1 - t_2}{2\delta(e)}$.

Profits are independent of taxes, and the total profit is normalized to 1. Note that, as in standard models, the amount profit shifting $\pi_i - \tilde{\pi}_i$ is proportional to the tax difference $t_i - t_j$ and inversely proportional to the enforcement level $\delta(e)$. Also, the tax base elasticity (sensitivity of the tax base to fiscal rates) is higher in the small country,[14] which is consistent with standard approaches.

[13] The firms' production decisions are $q_i^a = q_i^b = \gamma_i/(3\beta)$, $p_i - c_i = \gamma_i/3$, so that $\pi_i^a = \pi_i^b = \gamma_i^2/(9\beta)$. Based on this, the first-order condition for $\tilde{\pi}_i^k$ ($k = a,\ b,\ i = 1,\ 2$) becomes $-t_i + t_j - \partial C\left(\pi_i^k, \tilde{\pi}_i^k\right)/\partial\tilde{\pi}_i^k = 0$. They derive Lemma 1.

[14] Evaluated at equal tax rates, the tax base elasticities at $t_1 = t_2 = t$ are $-\dfrac{\partial\tilde{\pi}_1}{\partial t_1}\dfrac{t_1}{\tilde{\pi}_1} =$

## 2.3 Tax and Enforcement Timing

From Lemma 1, we have $\tilde{\pi}_i = \tilde{\pi}_i(t, e)$ and $R_i(t, e) = t_i\tilde{\pi}_i(t, e)$. The net tax revenue (2), taking account of the firms' behavior, is:

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

where $\epsilon_1 = \epsilon = -\epsilon_2$. Before we analyze the noncooperative choices of taxes and enforcement efforts, in this subsection, we discuss the issue of tax and enforcement timing.

We first 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. Then, when $\delta(e) = 0.5e_1 + 0.5e_2$, there is no equilibrium.*[15]

*Proof:* See the Appendix.

An intuition of Lemma 2 is as follows. If $t_i$ and $e_i$ are chosen simultaneously, from the point that satisfies the local first-order conditions, there always exist deviations to increase the tax revenue whereby managing the tax leak due to the profit shifting (mathematically, the Hessian matrix associated with the fist-order conditions does not satisfy negative definiteness).

For the rest of this paper, we adopt the following sequentiality of enforcement and tax decisions: (i) First, both countries set their enforcement efforts. (ii) Second, both countries choose their tax rates. The reasons we adopt this setup are as follows. First, the level of enforcement efforts 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 more easily, so they are assumed to be chosen before taxes (see Bacchetta and Espinosa (1995), Peralta et al. (2006) and Keen and Konrad (2013)).[16] [17] Second, by treating tax-enforcement decisions as a commitment device,

---

$\dfrac{t}{\delta(e)(1+\epsilon)}$ and $-\dfrac{\partial\tilde{\pi}_2}{\partial t_2}\dfrac{t_2}{\tilde{\pi}_2} = \dfrac{t}{\delta(e)(1-\epsilon)}$ .

[15] The enforcement function in Lemma 2 is used in our benchmark analysis.

[16] See also Bucovetsky and Haufler (2008) where governments decide the degree of tax preferences in favor of mobile capital before they set taxes on capital. The choice of such variables determines long-run reputation, based on which the governments decide the tax rates.

[17] When the sequence of the decisions is reversed so that taxes are chosen first and the enforcement is chosen later, then: (i) the subgame-perfect equilibrium exists only

this structure allows us to examine a practical issue of the possibility of enforcement cooperation under international tax competition. As to the tax timing, we first assume the conventional framework where countries choose tax rates simultaneously. As an extension, we examine sequential tax choices à la Hindriks and Nishimura (2015) in Section 4.2.

# 3 Tax and Enforcement Choices

## 3.1 Tax choices

The modified sequence of events is as follows. First, both countries set their enforcement efforts. Second, they compete in taxes. Third, MNEs choose a level of production and the profit to be shifted. The model is solved by backward induction. Hence, given the firms' decisions characterized in Lemma 1, in this section, we derive the governments' choices of taxes and enforcement efforts. As to the enforcement efforts, we compare the equilibria of the noncooperative and cooperative choice of enforcement to see whether the latter scenario is preferred unanimously or not in Section 3.3.

In the second stage, given the enforcement $e = (e_i, e_j)$, each country noncooperatively chooses its own tax rate $t_i$ $(i = 1, 2)$ to maximize the tax revenue:[18]

$$R_i(t_i, t_j, e) = t_i \tilde{\pi}_i(t_i, t_j, e) = t_i \left( \frac{1 + \epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)} \right),$$

where $\epsilon_1 = \epsilon = -\epsilon_2$. The first-order conditions are:

$$\frac{\partial W_i}{\partial t_i} = \frac{1 + \epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)} + t_i \frac{-1}{2\delta(e)} = 0. \tag{3}$$

The second-order conditions are satisfied. They yield the following equilibrium taxes:

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

As discussed in Hindriks and Nishimura (2015, p.66), the large country has greater market power.[19] In turn, the large country taxes higher with greater

if $\epsilon \geq 0.2$, (ii) only country 1 exerts enforcement efforts, and (iii) both the equilibrium taxes and enforcement efforts are lower than those obtained in Sections 3.1 and 3.2. The derivation is available upon request to the authors.

[18] As such, $e_i$ in (2) is treated as a sunk cost when countries choose taxes.

[19] This is an analogy to Amir and Stepanova (2006) in which the firms' cost asymmetry implies different markups.

asymmetry: $t_1^N(e) - t_2^N(e) = 2\dfrac{\delta(e)\epsilon}{3} > 0$. From Lemma 1, both firms shift profits from the large to the small (low-tax) country: we have $\tilde{\pi}_1^N = \dfrac{1+\epsilon}{2} - \dfrac{\epsilon}{3} = \dfrac{3+\epsilon}{6}$ and $\tilde{\pi}_2^N = \dfrac{3-\epsilon}{6}$. The tax revenues $R_i^N(e) = t_i^N(e)\tilde{\pi}_i^N$ $(i = 1, 2)$ are:

$$R_1^N(e) = \frac{\delta(e)}{2}\left(\frac{3+\epsilon}{3}\right)^2 \geq R_2^N(e) = \frac{\delta(e)}{2}\left(\frac{3-\epsilon}{3}\right)^2 \text{ for all } \epsilon \geq 0. \tag{4}$$

Note that, given $e$, total revenue is increasing in the extent of country asymmetry.

## 3.2 Noncooperative enforcement choices

In the first stage, taking account of the behavior in the subsequent stages, the government in each country chooses enforcement efforts, either noncooperatively or cooperatively. In this subsection, we first examine noncooperative (decentralized) enforcement choices.

As a benchmark, we consider the case of perfect substitutability between enforcement choices. Total enforcement level is given by

$$\delta(e) = 0.5e_1 + 0.5e_2. \tag{5}$$

Enforcement efforts increase the cost of profit shifting which in turn influences the choice of taxes. So country $i$, given $e_j$, maximizes $W_i^N(e_i, e_j) = R_i^N(e_i, e_j) - \eta\frac{e_i^2}{2}$ where $R_i^N(e)$'s are given in (4). The first-order conditions yield the equilibrium enforcement $e^N = (e_1^N, e_2^N)$ given by:[20]

**Lemma 3** $e_1^N = \dfrac{1}{\eta}\left(\dfrac{3+\epsilon}{6}\right)^2$, $e_2^N = \dfrac{1}{\eta}\left(\dfrac{3-\epsilon}{6}\right)^2$, $\delta(e^N) \equiv \delta^N = \dfrac{9+\epsilon^2}{36\eta}$. *$\delta^N$ is increasing in $\epsilon$: higher asymmetry induces higher enforcement in equilibrium.*

[20]See also (8) below. A general treatment is as follows:

$$t_i\left(\frac{t_i - t_j}{2(\delta(e))^2}\right)\frac{\partial\delta(e_i, e_j)}{\partial e_i} + \frac{\partial R_i^N(t_i, t_j, e)}{\partial t_j}\frac{\partial t_j^N}{\partial e_i} - \eta e_i = 0.$$

The first two terms capture the marginal benefit of the enforcement. When taxes and enforcement efforts are decided simultaneously, as shown in (11) in the Appendix, the low-tax country never has an incentive for enforcement. In contrast, the second term appears under our sequential structure, which gives an incentive to a more strict enforcement.

Lemma 3 illustrates the synergy between market power and enforcement. From (4), the marginal benefit of enforcement is proportional to tax revenue. With higher asymmetry, country 1 exhibits higher equilibrium taxes. In turn, the country is playing for high stakes in the enforcement game, which causes higher equilibrium level of enforcement: $e_1^N > e_2^N$, and $e_1^N$ increases quadratically in $\epsilon$. This increase of $e_1^N$ in $\epsilon$ is dominant in $\delta^N$ so that is $\delta^N$ is also increasing in $\epsilon$. This effect is similar to that of the total revenue noted above. We will come back to this point in Section 4.1 under enforcement complementarity.

Plugging the equilibrium enforcement into the welfare function, we have $W_i(e_1^N, e_2^N) = R_i^N(e_1^N, e_2^N) - \eta \dfrac{(e_i^N)^2}{2}$ $(i = 1, 2)$.

## 3.3 Benefit of enforcement cooperation

We now examine cooperative enforcement choices in the first stage, and see whether the cooperative framework is adopted unanimously or not. Here, both countries choose their enforcement levels so as to maximize their joint welfare. This reflects an agreement for the level of information exchange in the tax treaty. However, reflecting a real-world situation that most tax treaties do not have cooperative or harmonized tax rates, we assume that countries still decide 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 (4). That is,

$$\max_{e_i, e_j} \sum_i \left( R_i^N(e_i, e_j) - \eta \frac{e_i^2}{2} \right) = \max_{e_i, e_j} \ \delta(e_i, e_j) \cdot \left( \frac{(3+\epsilon)^2}{18} + \frac{(3-\epsilon)^2}{18} \right) - \eta \frac{e_i^2}{2} - \eta \frac{e_j^2}{2}. \tag{6}$$

The first-order conditions derive cooperative effort levels $e^{N*} = (e_1^{N*}, e_2^{N*})$ given by:

$$e_1^{N*} = e_2^{N*} = \frac{9 + \epsilon^2}{18\eta}$$

Enforcement efficiency requires both countries to exert the same enforcement effort, due to the quadratic (convex) cost function. Also, cooperative enforcement is increasing in the market asymmetry. 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. Enforcement cooperation induces the welfare levels $W_i(e_1^{N*}, e_2^{N*}) = R_i^N(e_1^{N*}, e_2^{N*}) - \eta \dfrac{(e_i^{N*})^2}{2}$ $(i = 1, 2)$. As to the benefit of enforcement cooperation, the large country unambiguously gains by cooperation ($W_1(e_1^{N*}, e_2^{N*}) > W_1(e_1^N, e_2^N)$ for all $\epsilon > 0$), due to higher enforcement for internalization of the enforcement externality. However, for country 2 there are offsetting effects. As $e_2^{N*}$ increases in $\epsilon$ but $e_2^N$ decreases in $\epsilon$ (Lemma 3), the cost of coordinated enforcement increases in $\epsilon$. As a result, even though the tax revenue increases ($R_2^N(e^{N*}) - R_2^N(e^N) > 0$), country 2 may prefer the noncooperative regime. The following proposition tells us that this is true under sufficient asymmetry.

**Proposition 1** *For the large country, $W_1(e_1^{N*}, e_2^{N*}) > W_1(e_1^N, e_2^N)$ for all parameter values. For the small country, on the other hand, $W_2(e_1^{N*}, e_2^{N*}) > W_2(e_1^N, e_2^N)$ iff $\epsilon < \epsilon^N \approx 0.2592817$. Therefore the small country prefers noncooperative enforcement when $\epsilon > \epsilon^N$.*

*Proof:* See the Appendix.

In the following section, we consider two different forces that can either promote or discourage enforcement efforts and cooperation. The first is related to the form of the enforcement technology (enforcement complementarity). The second is related to the form of tax competition (tax leadership).

# 4 Drivers of cooperation

## 4.1 Enforcement complementarity

Enforcement efforts increase the cost of profit shifting. However, in reality, dispersed (unilateral) enforcement efforts among the involved countries are not effective. For instance, the lack of tax-relevant information by host country makes the taxable income unclear to the home country. To formalize the notion of imperfect substitutability of enforcement efforts, we modify the enforcement function of (5) in the benchmark case. As an extension, we take the following CES formula:

$$\delta(e_1, e_2) = \left(0.5\, {e_1}^{-\rho} + 0.5\, {e_2}^{-\rho}\right)^{-\frac{1}{\rho}}, \quad \rho \geq -1. \tag{7}$$

The representative cases are: (i) $\rho = -1$ (perfect substitutes in (5)); (ii) $\rho \to 0$ (Cobb-Douglas case where $\delta(e_1, e_2) = e_1^{0.5} e_2^{0.5}$); and (iii) $\rho \to \infty$ (the weakest-link case where $\delta(e_1, e_2) = \min[e_1, e_2]$). If for example, the low-tax country can exercise the veto power during the mutual agreement procedure on the transfer price and taxable incomes of the MNEs, then the composition function becomes closer to the weakest-link formula.

The equilibrium strategy of the second stage (tax choice) is in fact the same as Section 3.1. That is, given $\delta(e)$, we obtain $R_i = R_i^N(e)$ $(i = 1, 2)$

in (4) as the equilibrium tax revenue. Recall that $R_1^N(e) \geq R_2^N(e)$ with the strict inequality when $\epsilon > 0$.

We then move on to examine the noncooperative choice of $e_i$ and $e_j$ in the first stage, when $\delta(e_1, e_2)$ takes the form of (7). Country $i$, given $e_j$, maximizes $W_i^N(e_i, e_j) = R_i^N(e_i, e_j) - \eta \frac{e_i^2}{2}$:

$$\frac{\partial \delta(e)}{\partial e_i} \frac{1}{2} \left( \frac{3+\epsilon_i}{3} \right)^2 - \eta e_i = 0 \quad \Rightarrow \quad \frac{1}{\eta} e_i^{-\rho-1} \left( 0.5 e_1^{-\rho} + 0.5 e_2^{-\rho} \right)^{\frac{1+\rho}{-\rho}} \left( \frac{3+\epsilon_i}{6} \right)^2 = e_i, \quad (8)$$

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

(8) constitutes the enforcement reaction function of country $i$. For $\rho = -1$, (8) is $\frac{1}{\eta} \left( \frac{3+\epsilon_i}{6} \right)^2 = e_i$. Namely, the reaction functions are orthogonal at $e_i = e_i^N$ in Lemma 3 as a dominant strategy. However, for $\rho > -1$, the enforcement reaction functions are upward-sloping, i.e., there is strategic complementarity in enforcement efforts. For the Cobb-Douglas case ($\rho = 0$), for example, (8) is $\frac{1}{\eta} e_j^{0.5} \left( \frac{3+\epsilon_i}{6} \right)^2 = e_i^{1.5}$. Figure 1 illustrates the cases of ($\rho = -0.2,\ \epsilon = 0.25$) and ($\rho = 5,\ \epsilon = 0.25$).

Figure 1 around here.

From (8), we obtain the following:

**Lemma 4** *$e_1^N(\rho) = \frac{1}{36\eta} \left( 0.5\ (\epsilon+3)^{\frac{-2\rho}{2+\rho}} + 0.5\ (3-\epsilon)^{\frac{-2\rho}{2+\rho}} \right)^{\frac{1+\rho}{-\rho}} (\epsilon+3)^{\frac{2}{2+\rho}}$, $e_2^N(\rho) = \frac{1}{36\eta} \left( 0.5\ (\epsilon+3)^{\frac{-2\rho}{2+\rho}} + 0.5\ (3-\epsilon)^{\frac{-2\rho}{2+\rho}} \right)^{\frac{1+\rho}{-\rho}} (3-\epsilon)^{\frac{2}{2+\rho}}$, and $\delta^N(\rho) = \frac{1}{36\eta} \left( 0.5\ (\epsilon+3)^{\frac{-2\rho}{2+\rho}} + 0.5\ (3-\epsilon)^{\frac{-2\rho}{2+\rho}} \right)^{\frac{2+\rho}{-\rho}}$.*

*$e_1^N(-1) > e_1^N(\rho) > e_2^N(\rho) > e_2^N(-1)$ for all $\rho > -1$. $\delta^N(\rho)$ is decreasing in $\rho$ for all $\rho \geq -1$. $\delta^N(\rho)$ is increasing in $\epsilon$ when $\rho < -\frac{2}{3}$, and $\delta^N(\rho)$ is decreasing in $\epsilon$ when $\rho > -\frac{2}{3}$.*

*Proof:* See the Appendix.

In Figure 1, the dotted curve **WP** illustrates the trajectory of $(e_1^N(\rho), e_2^N(\rho))$ on $(e_1, e_2)$-space (with $\epsilon = 0.25$). It starts with **P**$=(e_1^N(-1), e_2^N(-1))$. Given $e_2 < e_1$, greater complementarity (larger $\rho$) reduces the marginal benefit of country 1 (LHS of (8)) and increases that of country 2. As a result, $e_1^N(\rho)$

decreases in $\rho$, whereas $e_2^N(\rho)$ increases in $\rho$. As $\rho$ increases, $(e_1^N(\rho), e_2^N(\rho))$ converges to $\mathbf{W}$=$(e_1^N(\infty), e_2^N(\infty))$, where:

$$e_i^N(\infty) = \frac{(3-\epsilon)^2}{36\eta}\frac{(3+\epsilon)^2}{\epsilon^2+9}$$

Figure 2 around here.

Figure 2 illustrates $e_1^N(\rho)$ and $e_2^N(\rho)$ for different $\epsilon$'s. Given $\rho$, when asymmetry increases, there are two effects for the determination of $e_1^N(\rho)$. First, asymmetry per se tends to increase $e_1$. Second, on the other hand, the decrease of $e_2^N(\rho)$ tends to reduce $e_1$, due to strategic complementarity. From these effects, $e_1^N(\rho)$ at $\epsilon = 0.25$ (the dashed curve located above) single-crosses $e_1^N(\rho)$ at $\epsilon = 0.1$ (the solid curve located above) from above: there is a critical value of $\rho$ above (below) which higher asymmetry induces lower (higher) enforcement by country 1 in equilibrium. In this way, there is interaction between asymmetry and complementarity. As to the enforcement level, higher asymmetry induces higher equilibrium enforcement as in Lemma 3 if and only if the technology is sufficiently substitutable, whereas higher $\epsilon$ induces lower $\delta^N(\rho)$ when $\rho$ is sufficiently high.

Next, cooperative effort levels are given by $\max_{e_i,e_j}\sum_i\left(R_i^N(e_i,e_j)-\eta\frac{e_i^2}{2}\right)$, which is the same as (6) of Section 3.3. With the enforcement function (7), the first-order conditions are now $\left(0.5e_1^{-\rho}+0.5e_2^{-\rho}\right)^{\frac{1+\rho}{-\rho}}\left(\frac{9+\epsilon^2}{18\eta}\right) = e_i^{2+\rho}$ $(i = 1,2)$. Invariant with respect to imperfect substitutability, we obtain $e_1^{N*} = e_2^{N*} = \frac{9+\epsilon^2}{18\eta}$ as in the basic model. With greater complementarity, noncooperative solution becomes less attractive, as $\delta^N(\rho)$ decreases in $\rho$. Recall that, when $\rho = -1$, the large country unambiguously gains by cooperation ($W_1(e_1^{N*}, e_2^{N*}) > W_1(e_1^N, e_2^N)$ for all $\epsilon > 0$) but it is not the case for the small country: $W_2(e_1^{N*}, e_2^{N*}) > W_2(e_1^N, e_2^N)$ if and only if $\epsilon < \epsilon^N \approx 0.2592817$ (Proposition 1). In comparison to the basic case, when $\rho > -1$, we show the following:

**Proposition 2** *Suppose that the enforcement function is represented by (7) with $\rho > -1$. For the large country, $W_1(e_1^{N*}(\rho), e_2^{N*}(\rho)) > W_1(e_1^N(\rho), e_2^N(\rho))$ for all parameter values. For the small country, on the other hand, $W_2(e_1^{N*}(\rho), e_2^{N*}(\rho)) > W_2(e_1^N(\rho), e_2^N(\rho))$ iff $\epsilon < \epsilon^N(\rho)$. $\epsilon^N(\rho)$ is an increasing function of $\rho$, with $\epsilon^N(-1) = \epsilon^N \approx 0.2592817$ in Proposition 1 and $\epsilon^N(\infty) \approx 0.4094092130 < 1$.*

*Proof:* See the Appendix.

Figure 3 around here.

For given level of complementarity ($\rho$), Figure 2 above tells us two effects of the increase in asymmetry ($\epsilon$). First, the gap of $e_1^N(\rho)$ and $e_2^N(\rho)$ tends to increase. Second, there is an ambiguity of the increase of country 1's enforcement effort (as well, the free-riding benefit of country 2) with respect to an increase in asymmetry. Proposition 2 tells us that the first effect is crucial for the benefit of cooperation by country 2, in that country 2 prefers the noncooperative regime under sufficient asymmetry: as discussed above in Proposition 1, a driving force of country 2's deviation is the conflict of interest with respect to the desired level of the enforcement. However, cooperation is more viable with greater enforcement complementarity: for example, $\epsilon^N(-0.2) \approx 0.3114695738$, $\epsilon^N(0) \approx 0.319521$, and $\epsilon^N(\infty) \approx 0.4094092130$. This is because the difference in $e_1^N(\rho)$ and $e_2^N(\rho)$ decreases as $\rho$ increases. On $(\rho, \epsilon)$-space, Figure 3 divides the area where cooperation is viable (below/right of the curve $\epsilon^N(\rho)$). Note, however, that even with perfect complementarity ($\rho = \infty$ at which $e_1^N(\rho) = e_2^N(\rho)$), we have $\epsilon^N(\infty) < 1$: with sufficient asymmetry, country 2 prefers the noncooperative regime with lower enforcement efforts.

## 4.2 Tax leadership

In reality, even if the countries do not reach in cooperation in tax decisions, the countries can enhance welfare within a noncooperative framework by adopting tax leadership (namely, Stackelberg tax competition). Following Hindriks and Nishimura (2015), we consider the case in which (large) country 1 leads. At the end of this subsection, we briefly discuss the case of the small country's leadership.

Keeping our sequential structure of enforcement and tax choices, we solve the game by backward induction. Given the enforcement choices $e = (e_i, e_j)$, the (small) country 2, as the tax follower, chooses $t_2$, given $t_1$. Country 1, as the Stackelberg leader, maximizes $R_1(t_1, t_2, e)$ with respect to $t_1$, along the country 2's tax reaction function $t_2 = \hat{t}_2(t_1; e)$. This gives the equilibrium tax rates denoted by $(t_1^S(e), t_2^S(e))$ and the tax revenues $R_i^S(e) = t_i^S(e)\tilde{\pi}_i^S$ ($i = 1, 2$) as follows (see the proof of Lamma 5 for derivation):

$$t_1^S(e) = \delta(e)\left(\frac{3+\epsilon}{2}\right), \; t_2^S(e) = \delta(e)\left(\frac{5-\epsilon}{4}\right), \; R_1^S(e) = \frac{\delta(e)}{16}(3+\epsilon)^2, \; R_2^S(e) = \frac{\delta(e)}{32}(5-\epsilon)^2. \tag{9}$$

Several explanations are in order. First, with the given level of $e$, $t_i^S(e) > t_i^N(e)$ $(i = 1, 2)$, due to the strategic complementarity in tax choices $(\partial \hat{t_2}(t_1; e)/\partial t_1 > 0)$. Second, the tax differential is $t_1^S(e) - t_2^S(e) > t_1^N(e) - t_2^N(e)$. And third, we have:

$$R_1^S(e) \gtreqless R_2^S(e) \iff \epsilon \gtreqless \hat{\epsilon} \approx 0.31370850.$$

Namely, when $\epsilon$ is small enough, the benefit of the second-mover advantage in the tax competition dominates the initial exogenous asymmetry, so that the tax leader collects less tax revenue than the tax follower $(R_1^S(e) < R_2^S(e))$, and vice versa when $\epsilon$ is large enough. As discussed later, this feature affects the incentive for tax enforcement.

We then move backward to compute noncooperative enforcement choices, followed by Stackerlberg taxation, $e^S = (e_1^S, e_2^S)$. As to the enforcement function, in this subsection we go back to the benchmark formulation of $\delta(e_1, e_2) = (0.5\, e_1 + 0.5\, e_2)$ (equation (5)). Indeed, the cost of profit shifting, measured by $\delta(e)$, affects equilibrium tax levels: higher $\delta(e)$ lowers tax competition and increases equilibrium taxes. The question is how the enforcement efforts will be changed under tax leadership.

**Lemma 5** *(i)* $e_1^S = \dfrac{(3+\epsilon)^2}{32\eta} > e_1^N,\ e_2^S = \dfrac{(5-\epsilon)^2}{64\eta} > e_2^N,\ \delta(e_1^S, e_2^S) \equiv \delta^S > \delta^N.$ *Namely, 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)* $|e_1^S - e_2^S| \lesseqgtr e_1^N - e_2^N \iff \epsilon \gtreqless \tilde{\epsilon} \approx 0.16094072 < \hat{\epsilon}$. *Namely, the Stackelberg tax competition reduces the dispersion of the enforcement efforts, provided that asymmetry is sufficiently large.*

*Proof:* See the Appendix.

Both countries get higher revenue in the Stackelberg subgame given the same effort as in the Nash subgame ($R_i^S(e) > R_i^N(e)$ $(i = 1, 2)$ from (4) and (9)). As a result, the marginal benefit of effort is higher in the Stackelberg situation ($e_i^S > e_i^N$ $(i = 1, 2)$ in part (i)). Endogenous enforcement further increases the relative benefit of the Stackelberg outcome $(R_i^S(e^S) > R_i^S(e^N) > R_i^N(e^N)$ $(i = 1, 2))$. Also, from $\delta^S > \delta^N$ in part (i) and (9), the tax gap is also increased: $t_1^S(e^S) - t_2^S(e^S) > t_1^S(e^N) - t_2^S(e^N) > t_1^N(e^N) - t_2^N(e^N)$. Similar to a "minimum tax" over the Nash (simultaneous) tax competition in Kanbur and Keen (1993), the benefit by the large country comes from the higher tax revenue along the tax reaction function, and at the same time, the small country benefits from the higher tax difference, due to $\partial \hat{t_2}(t_1; e)/\partial t_1 < 1$, that allows more profit shifting in our context.[21] As emphasized in Kanbur and Keen (1993), this type of coordination strategy is different from tax harmonization that eliminates the benefit of the profit shifting to the small country.

The reason for part (iii) is as follows. In contrast with the Nash tax competition where $e_1^N > e_2^N$ for all parameter values, the large country may exert *lower* enforcement effort if the asymmetry is small enough: from part (i) we have $e_1^S \gtrless e_2^S \iff \epsilon \gtrless \hat{\epsilon}$. This means that the dispersion of enforcement efforts is lower under Stackelberg competition than under Nash tax competition if $\epsilon$ is close to $\hat{\epsilon}$, and in fact, the range of asymmetries where $|e_1^S - e_2^S| < e_1^N - e_2^N$ holds is quite large.

We next 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 $e = (e_1, e_2)$ so as to maximize their joint welfare but countries compete in taxes subsequently. Therefore, given $e$, countries choose $(t_1^S(e), t_2^S(e))$. The choice of $e_i$ and $e_j$ is given by $\max_{e_i,e_j} \sum_i \left( R_i^S(e_i, e_j) - \eta \frac{e_i^2}{2} \right)$. In the Appendix we show the following:

**Proposition 3** *Under the Stackelberg tax competition, the enforcement cooperation is more viable than under the Nash tax competition:* $W_i(e_1^{S*}, e_2^{S*}) > W_i(e_1^S, e_2^S)$ $(i = 1, 2)$ *for* $\epsilon < \epsilon^S \approx 0.6542911$ *with* $\epsilon^S > \epsilon^N$.

*Proof:* See the Appendix.

The tax leadership (by the large country) facilitates the participation of the small country in the enforcement cooperation. There are two reasons. First, $R_i^S(e^S) > R_i^N(e^N)$ $(i = 1, 2)$ as in Hindriks and Nishimura (2015), so that the benefit of internalization of enforcement externality is larger. Second, the enforcement gap is reduced as we showed in Lemma 5. (iii). As in our previous propositions, the second effect reduces the conflict of interest with respect to the desired level of the enforcement.

Lastly, we discuss the issue of the leadership.[22] Let $R_i^M(e)$ $(i = 1, 2)$ be the country $i$'s revenue under country 2's leadership with given level of

[21]The difference between the minimum tax and the large country's leadership is that the former entails a move along the large country's reaction function, whereas we consider here the move along the small country's reaction function.

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

$e$. Then we have $R_1^M(e) > R_1^S(e) > R_1^N(e)$ and $R_2^S(e) > R_2^M(e) > R_2^N(e)$. Namely, country 2's leadership also increases countries' enforcement incentives compared with the Nash tax competition. However, regarding 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 $\epsilon$ (enforcement gap is increased by the leadership). As a result, for the cooperative effort level $(e_1^{M*}, e_2^{M*})$, $W_2(e_1^{M*}, e_2^{M*}) < W_2(e_1^M, e_2^M)$ if and only if $\epsilon > \epsilon^M \approx 0.01692095340$. That is, the threshold value of viable cooperation is even smaller than $\epsilon^N$. Moreover, we can show that the large leadership equilibrium Pareto dominates the small leadership equilibrium under sufficient asymmetry, as in Hindriks and Nishimura (2015).[23] According to Proposition 3, the former equilibrium leads to higher tax revenues to both countries as well as more viable enforcement coordination than under the Nash tax competition.

# 5 Conclusion

Using a profit shifting model with multinational enterprises that operate in two countries with different market sizes, we analyze the determinants of voluntary cooperation on enforcement effort and information sharing arrangements. When the enforcement efforts exhibit complementarity, we showed that cooperation is more viable with greater enforcement complementarity (less additive/substitutable nature). The exact specification comes from administrative or legal nature of the collective decision-making. However, we can infer in which situations the enforcement coordination (by unanimity) is more viable; namely, in the situation where the capital-importing country can exercise a veto power towards tighter enforcement, and where bilateral (less dispersed) efforts by host and home countries are required for effective information exchange.

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 more enforcement efforts than under the simultaneous tax choice, and the enforcement cooperation will be more viable under the Stackelberg tax competition. Interestingly, this type of coordination strategy is different from tax harmonization that eliminates the benefit of the profit shifting to the small country.

The action programs of the OECD BEPS Project have various degree of commitment, from the strong commitment to consistent implementation of the program across countries to the weak form where countries are free to

[23] $W_2(e_1^M, e_2^M) < W_2(e_1^S, e_2^S)$ for all $\epsilon$, and when $\epsilon > 0.6618315754$, $W_1(e_1^M, e_2^M) < W_1(e_1^S, e_2^S)$.

assess and pick. Looking forward, this paper has two contributions. First, developing a simple model that captures the central features of profit shifting, we produced sharp insights into the costs and benefits of enforcement cooperation. Second, given the rapid development of the empirical literature, the analysis here guides us what empirical quantities to look for, towards a practical prediction of the viability of the enforcement cooperation.

# Appendix

**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:

$$\frac{\partial W_i}{\partial t_i} = \frac{1+\epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)} + t_i \frac{-1}{2\delta(e)} = 0, \tag{10}$$

$$\frac{\partial W_i}{\partial e_i} = t_i \left( \frac{t_i - t_j}{2(\delta(e))^2} \right) \frac{\partial \delta(e_i, e_j)}{\partial e_i} - \eta e_i \leq 0. \tag{11}$$

(10) derives the preferred tax level by 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 $\partial\delta(e_i, e_j)/\partial e_i > 0$, (11) implies that country 2 chooses $e_2 = 0$. When $\delta(e) = 0.5e_1 + 0.5e_2$, from (10) and (11), the set of values that satisfy the first-order necessary condition is uniquely given by $e_1^n = \frac{\epsilon\,(3+\epsilon)}{18\eta}$, $e_2^n = 0$ and $t_1^n = 0.5e_1^n \left(\frac{3+\epsilon}{3}\right)$ and $t_2^n = 0.5e_1^n \left(\frac{3-\epsilon}{3}\right)$. However, evaluated at $(e_2^n, t_2^n)$, the Hessian matrix of country 1 associated with (10) and (11) does not satisfy negative definiteness: $\begin{vmatrix} \frac{\partial^2 W_1}{(\partial t_1)^2} & \frac{\partial^2 W_1}{\partial t_1 \partial e_1} \\ \frac{\partial^2 W_1}{\partial e_1 \partial t_1} & \frac{\partial^2 W_1}{(\partial e_1)^2} \end{vmatrix} = \frac{-2\epsilon^2(3-\epsilon)^2(3+\epsilon)^2 + (36\eta e_1)^3}{23328\eta^2 e_1^4}$. For $e_1^a = \frac{\epsilon\,(3-\epsilon)}{36\eta}$ ($e_1^a < e_1^n$) and $t_1^a = \frac{\epsilon(5\epsilon+9)(3-\epsilon)}{432\eta}$, we have $W_1(e_1^a, t_1^a, e_2^n, t_2^n) - W_1(e_1^n, t_1^n, e_2^n, t_2^n) = \frac{\epsilon\,(1+\epsilon)^2\,(1-\epsilon)}{192\eta}$, which is positive for $\epsilon > 0$. For $\epsilon = 0$, we set $e_1^b = \frac{1}{36\eta}$ ($e_1^b > e_1^n$) and $t_1^b = \frac{1}{144\eta}$, and we have $W_1(e_1^b, t_1^b, e_2^n, t_2^n) - W_1(e_1^n, t_1^n, e_2^n, t_2^n) = \frac{7}{5184\eta} > 0$. Therefore, the equilibrium does not exist. *Q.E.D.*

**Proof of Proposition 1:** $R_2^N(e^{N*}) - R_2^N(e^N) = \left(\frac{9+\epsilon^2}{18\eta}\right)\left(\frac{3-\epsilon}{6}\right)^2$ is decreasing in $\epsilon$, whereas $\eta\frac{(e_2^{N*})^2}{2} - \eta\frac{(e_2^N)^2}{2} = \left(\frac{3+\epsilon}{6}\right)^2 \left(\frac{9+\epsilon^2-2\epsilon}{24\eta}\right)$ is increasing in $\epsilon$. From these, $W_2^N(e^{N*}) - W_2^N(e^N) = \frac{81+\epsilon^4-36\epsilon^3+54\epsilon^2-324\epsilon}{2\cdot 36^2\eta}$, which is positive (negative) when $\epsilon < \epsilon^N$ ($\epsilon > \epsilon^N$). The analogous calculation shows $W_1^N(e^{N*}) - W_1^N(e^N) = \frac{81+\epsilon^4+36\epsilon^3+54\epsilon^2+324\epsilon}{2\cdot 36^2\eta} > 0$ for all $\epsilon$. *Q.E.D.*

**Proof of Lemma 4:** Solving the system of (8) for $i = 1, 2$, we obtain the expression of $(e_1^N(\rho), e_2^N(\rho))$ and $\delta^N(\rho)$. Clearly $e_1^N(\rho) > e_2^N(\rho)$. Let $f(x,y) \equiv \left(0.5\, x^{\frac{-2\rho}{2+\rho}} + 0.5\, y^{\frac{-2\rho}{2+\rho}}\right)^{\frac{1+\rho}{-\rho}}$. Since $f$ is increasing in both $x$ and $y$, one has $f(3+\epsilon, 3+\epsilon) > f(3+\epsilon, 3-\epsilon)$ and $f(3+\epsilon, 3-\epsilon) > f(3-\epsilon, 3-\epsilon)$, which is equivalent to $e_1^N(-1) > e_1^N(\rho)$ and $e_2^N(\rho) > e_2^N(-1)$ respectively.

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

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

from Jensen's inequality, since the function $g(x) = x \ln x$ is convex with respect to $x$.

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

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

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

**Proof of Proposition 2**: $W_2(e_1^{N*}, e_2^{N*}) = \left(\frac{9+\epsilon^2}{18\eta}\right)\left(\frac{9+\epsilon^2-12\epsilon}{36}\right)$ as in the benchmark case, whereas $W_2(e_1^N(\rho), e_2^N(\rho)) = \left(\frac{3-\epsilon}{6}\right)^2 \times$
$\left(0.5\ (\epsilon+3)^{\frac{-2\rho}{2+\rho}} + 0.5\ (3-\epsilon)^{\frac{-2\rho}{2+\rho}}\right)^{\frac{2+2\rho}{-\rho}} \frac{\left((3-\epsilon)^{\frac{-2\rho}{2+\rho}} + 2\ (3+\epsilon)^{\frac{-2\rho}{2+\rho}}\right)}{72\eta}$. From these, we numerically show that $W_2^N(e^{N*}) - W_2^N(e^N(\rho))$ is positive (negative) when

$\epsilon < \epsilon^N(\rho)$ ($\epsilon > \epsilon^N(\rho)$). $\epsilon^N(\rho)$ is shown in Figure 3. The analogous calculation shows $W_1^N(e^{N*}) - W_1^N(e^N(\rho)) > 0$ for all $\epsilon$. *Q.E.D.*

**Proof of Lemma 5**: From (3), the tax reaction function of country 2 is $\arg\max_{t_2} R_2(t_2, t_1, e) \equiv \hat{t_2}(t_1; e) = \delta(e)\left(\frac{1-\epsilon}{2}\right) + \frac{t_1}{2}$. We have $\frac{\partial \hat{t_2}}{\partial t_1} \in (0,1)$ as in conventional models of tax competition. The first-order condition of country 1 is given by $\frac{\partial R_1}{\partial t_1} + \frac{\partial R_1}{\partial t_2}\frac{\partial \hat{t_2}}{\partial t_1} = \frac{1+\epsilon}{2} - \frac{t_1 - \hat{t_2}(t_1)}{2\delta(e)} - t_1\frac{1-(1/2)}{2\delta(e)} = 0$, which yields $t_1 = t_1^S(e)$ and $t_2^S(e) = \hat{t_2}(t_1^S(e); e)$ in (9).[24] $t_1^S(e) - t_2^S(e) = \delta(e)\frac{1+3\epsilon}{4}$. From Lemma 1, $\tilde{\pi}_1^S = \frac{1+\epsilon}{2} - \frac{1+3\epsilon}{2\cdot 4} = \frac{3+\epsilon}{8} < \tilde{\pi}_2^S = \frac{5-\epsilon}{8}$. We then have $R_i^S(e) = t_i^S(e)\tilde{\pi}_i^S$ $(i = 1, 2)$ as in (9).

From (9) and (5) in the text, the maximization of $W_i^S(e_i, e_j) = R_i^S(e_i, e_j) - \eta\frac{e_i^2}{2}$ with respect to $e_i$ yields $e_1^S = \frac{(3+\epsilon)^2}{32\eta}$ and $e_2^S = \frac{(5-\epsilon)^2}{64\eta}$. $e_1^S - e_1^N = \frac{1}{288\eta}(3+\epsilon)^2 > 0$, $e_2^S - e_2^N = \frac{1}{576\eta}(3+\epsilon)(27-7\epsilon) > 0$. Therefore, $\delta^S > \delta^N$, so part (i) of the lemma holds. The proof of part (ii) of the lemma was shown in the text. For $\epsilon < \hat{\epsilon} \approx 0.31370850$, $0 < e_2^S - e_1^S = \frac{7-\epsilon^2-22\epsilon}{64\eta} < e_1^N - e_2^N = \frac{\epsilon}{3\eta}$ if and only if $\epsilon > \tilde{\epsilon} \approx 0.16094072$. For all $\epsilon \geq \hat{\epsilon}$, $0 \leq e_1^S - e_2^S = \frac{-7+\epsilon^2+22\epsilon}{64\eta} < e_1^N - e_2^N$. Therefore, part (iii) of the lemma holds. *Q.E.D.*

**Proof of Proposition 3**: The welfare levels under the noncooperative enforcement choice (followed by Stackelberg tax competition) are:

$$W_1(e_1^S, e_2^S) = \left(\frac{3+\epsilon}{4}\right)^2\left(\frac{17+\epsilon^2-2\epsilon}{64\eta}\right),\ W_2(e_1^S, e_2^S) = \left(\frac{5-\epsilon}{8}\right)^2\left(\frac{61+5\epsilon^2+14\epsilon}{128\eta}\right),$$

The joint welfare maximizing enforcement is

$$e_1^{S*} = \frac{3\epsilon^2+2\epsilon+43}{64\eta} = e_2^{S*}$$

The cooperative welfare levels, $W_i(e_1^{S*}, e_2^{S*}) = R_i^S(e_1^{S*}, e_2^{S*}) - \eta\frac{(e_i^{S*})^2}{2}$ $(i =$

[24] Note that the revenue function is concave in the tax rate, $\partial\left(\frac{\partial R_1}{\partial t_1} + \frac{\partial R_1}{\partial t_2}\frac{\partial \hat{t_2}}{\partial t_1}\right)/\partial t_1 = -1/(2\delta(e)) < 0$.

1, 2), are:

$$W_1(e_1^{S*}, e_2^{S*}) = \left(\frac{3\epsilon^2 + 2\epsilon + 43}{64\eta}\right)\left(\frac{29 + 5\epsilon^2 + 46\epsilon}{128}\right),$$
$$W_2(e_1^{S*}, e_2^{S*}) = \left(\frac{3\epsilon^2 + 2\epsilon + 43}{64\eta}\right)\left(\frac{\epsilon^2 - 42\epsilon + 57}{128}\right).$$

For the large country, $W_1^N(e^{S*}) - W_1^N(e^S) = \dfrac{23 + 7\epsilon^4 + 116\epsilon^3 + 282\epsilon^2 + 1364\epsilon}{2 \cdot 64^2\eta} >$ 0 for all $\epsilon$. For the small country, $W_2^N(e^{S*}) - W_2^N(e^S) = \dfrac{463 - \epsilon^4 - 44\epsilon^3 + 42\epsilon^2 - 716\epsilon}{64^2\eta}$, which is positive when $\epsilon < \epsilon^S \approx 0.6542911$, and negative when $\epsilon > \epsilon^S$. This critical value of asymmetry is greater than $\epsilon^N \approx 0.2592817$ in the Nash game. *Q.E.D.*

# References


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

[2] Amir, R. and A. Stepanova, 2006. Second-mover advantage and price leadership in Bertrand duopoly, *Games and Economic Behavior*, vol. 55, pages 1-20.

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

[4] Auerbach, A. M. Devereux, and H. Simpson, 2010. Taxing corporate income. In J. Mirrlees et al., eds., *Dimension of tax design: The Mirrlees review.* Oxford: Oxford University Press, pages 837-913.

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

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

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

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

[9] Clausing, K., 2003. Tax motivated transfer pricing and US intra-firm trade prices, *Journal of Public Economics*, vol. 87, pages 2207-2223.

[10] Cremer, H., and Gahvari, F., 2000. Tax evasion, fiscal competition and economic integration, *European Economic Review*, vol. 44, pages 1633-1657.

[11] Devereux, M. and R. Griffith, 2003. Evaluating tax policy for location decisions, *International Tax and Public Finance*, vol. 10, pages 107-126.

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

[13] Devereux, M., B. Lockwood, and M. Redoano, 2008. Do countries compete over corporate tax rates? *Journal of Public Economics*, vol. 92, pages 1210-1235.

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

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

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

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

[18] Hindriks, J., S. Peralta, and S. Weber, 2014. Local taxation of global corporation: a simple solution, *Annals of Economics and Statistics*, vol 113-114, pages 37-65.

[19] Hirshleifer, J., 1983. From weakest-link to best-shot: the voluntary provision of public goods, *Public Choice*, vol. 41, pages 371-386.

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

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

[22] IMF, 2011. *Appendix B: estimating bank responses to CIT rates.*

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

[24] Keen, M., and K. Konrad, 2013. The theory of international tax competition and coordination, In A. Auerbach et al., eds., *Handbook of Public Economics*, volume 5, North Holland, pages 257-328.

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

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

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

[28] Lohse, T. and Riedel, N. 2013. Do transfer pricing laws limit international income shifting? Evidence from european multinationals, *CESifo Working Paper*, 4404.

[29] Mintz, J. and M. Smart, 2003. "Income shifting, investment, and tax competition: theory and evidence from provincial taxation in Canada", *Journal of Public Economics*, vol. 88, pages 1149-1168.

[30] Næss-Schmidt, P., E. Sunesen, M. Hansen, and D. Ali, 2012. Tax elasticities of financial instruments, profits and remuneration, WORKING PAPER N.30-2012, European Union.

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

[32] OECD, 2015. *Base Erosion and Profit Shifting.* http://www.oecd.org/tax/beps.htm

[33] Peralta, S., X. Wauthy, and T. van Ypersele, Tanguy, 2006. Should countries control international profit shifting?, *Journal of International Economics*, vol. 68, pages 24-37.

[34] Redoano, M., 2007. Fiscal interactions among European countries. Does the EU matter? CESifo Working Paper Series No. 1952.

[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*, vol. 15, pages 185-207.

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

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

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

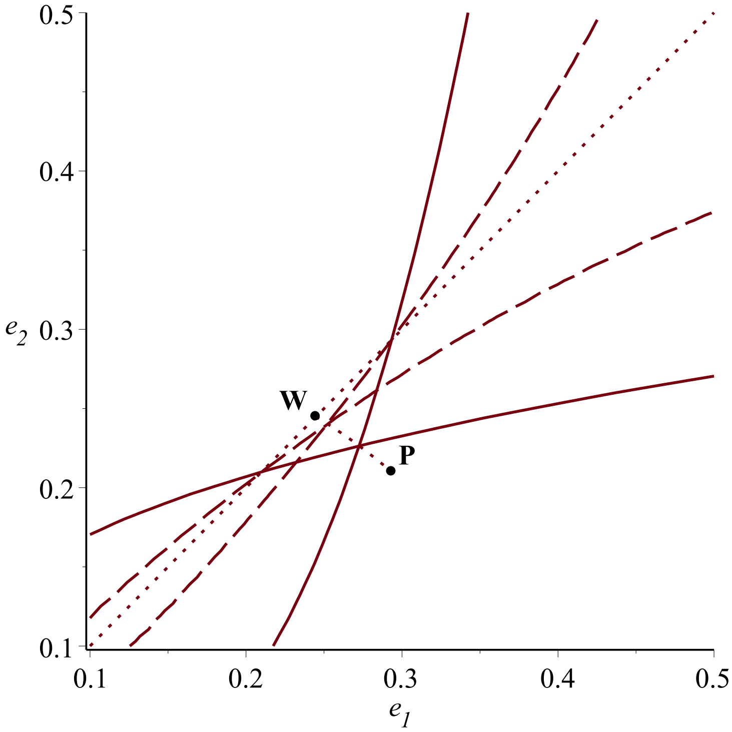


Figure 1: Enforcement reaction functions for $\rho = -0.2$ (solid curves) and $\rho = 5$ (dashed curves). The curve **WP** is the trajectory of $(e_1^N(\rho), e_2^N(\rho))$.

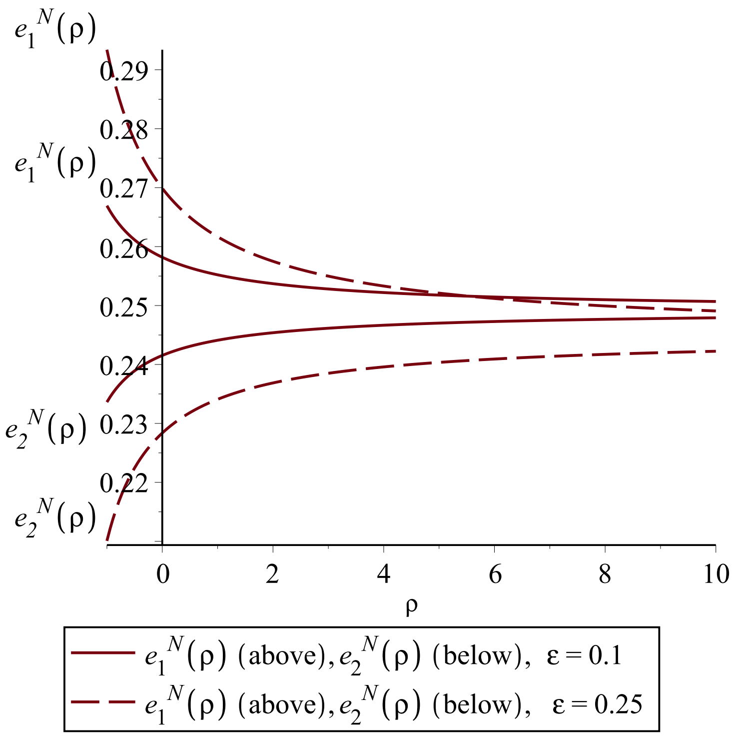


Figure 2: Noncooperative enforcement level for $\epsilon = 0.1$ and $\epsilon = 0.25$

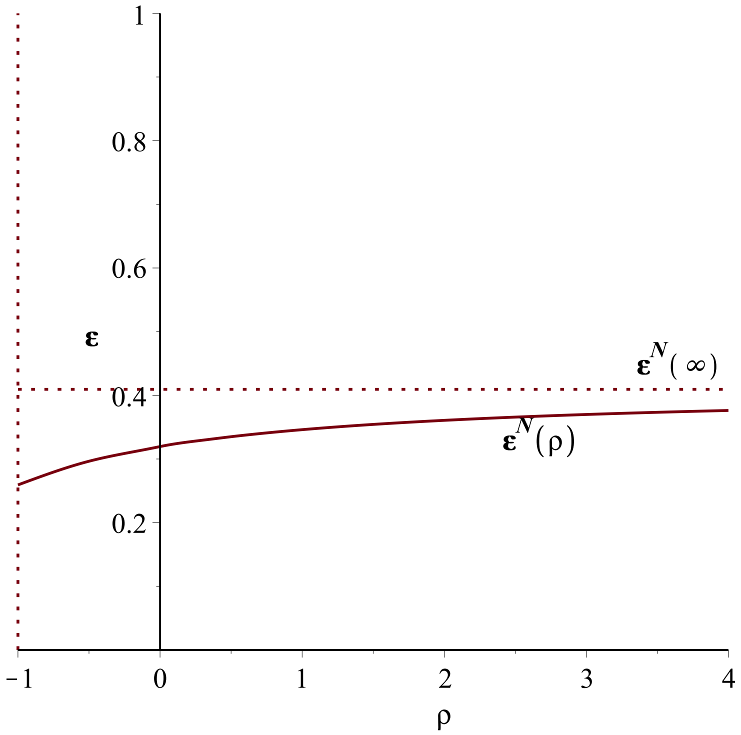


Figure 3: Critical value of asymmetry $\epsilon^N(\rho)$