Received: 12 December 2020 | Accepted: 27 December 2020
DOI: 10.1111/jpet.12501

ORIGINAL ARTICLE

JOURNAL OF PUBLIC ECONOMIC THEORY WILEY

Taxing multinationals: The scope for enforcement cooperation

Jean Hindriks[1] | **Yukihiro Nishimura**[2,3]

[1]CORE (LIDAM) and Economics School of Louvain, Catholic University of Louvain, Louvain-la-Neuve, Belgium
[2]Graduate School of Economics, Osaka University, Osaka, Japan
[3]CESifo, Munchen, Germany

Correspondence
Yukihiro Nishimura, Graduate School of Economics, Osaka University, 1-7 Machikaneyama-cho, Toyonaka-shi, Osaka, 560-0043, Japan.
Email: ynishimu@econ.osaka-u.ac.jp

Funding information
Japan Legislative Society Foundation; Obayashi Foundation; MEXT/JSPS KAKENHI, Grant/Award Numbers: JP15K03511, JP18H00866

Abstract
Policymakers seeking to raise more tax revenues from multinational enterprises have two alternatives: to raise tax rates or to devote more resources to improve tax compliance. Tougher tax enforcement increases the cost of profit shifting, and thus mitigates tax competition. We present a tax-competition model with two policy instruments (the corporate tax rate and the tightness of tax enforcement). In line with the Organisation for Economic Cooperation and Development's Base Erosion and Profit Shifting project, we analyze the scope for enforcement cooperation among asymmetric countries, considering that taxes are set noncooperatively. We show that the low-tax country may fail to cooperate if asymmetry is large enough and that tax havens would never agree to cooperate. Then we identify two drivers for enforcement cooperation. The first driver of cooperation is the complementarity of enforcement actions across countries. This is because the efficiency loss from enforcement dispersion is greater under complementarity. The second driver of cooperation is tax leadership by the high-tax country, which acts as a level-playing field in the tax competition and reduces the extent of disagreement on enforcement.

1 INTRODUCTION

New technologies and the globalization of the economy have facilitated tax avoidance through the shifting of profits by multinational enterprises (MNEs) to low or no-tax jurisdictions (tax havens). This is the essence of base erosion and profit shifting (BEPS, hereafter referred to as "profit shifting").[1] To address this issue, the international community has made substantial efforts, but uncoordinated actions across countries led to a patchwork of unilateral actions, which dampened their overall effectiveness for the governments to collect tax revenues. Facing this challenge, the Organisation for Economic Cooperation and Development's (OECD's) (2015) BEPS report proposed global action plans containing a series of multilateral enforcement efforts to be undertaken by governments. The European Union (EU) also established the Fiscalis 2020 program (European Commission, 2020) 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 result, which is the reduction of aggressive tax planning to mitigate the tax competition, affects the other country. Therefore, tax enforcement shares the properties of joint production with individual efforts, where the sharing of the production gains leads to a nontrivial incentive problem.[2] 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. This case reflects the current OECD framework to reinforce enforcement cooperation in which each country still can freely choose its tax system and tax rates.[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

[1]There is convincing empirical evidence of profit shifting. 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% of the additional revenue resulting from a unilateral tax increase is lost because of profit shifting. See also Swenson (2001), Clausing (2003, 2009), Huizinga and Laeven (2008), Næss-Schmidt et al. (2012), and Grubert and Altshuler (2013). See also Becker et al. (2020) on the allocation of risk.

[2]Holmstrom (1982) solved the incentive problem by penalizing the entire group when any member of the group shirks. In our international context, there is no central authority that can implement such a group punishment.

[3]Our tax enforcement cooperation is in contrast with the tax cooperation in Cremer and Gahvari (2000). They showed that tax harmonization could be damaging because it would induce countries to adopt lenient enforcement policies. There is no tax cooperation in our model because the OECD BEPS project is rather to encourage enforcement cooperation letting each country freely choose their taxes.

application of the OECD's (2015) proposals to the G20 and the EU because the adoption of such proposals often requires unanimity.

In this context, we identify two different remedies for the lack of cooperation. The first remedy 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. Indeed, recommendations made in the OECD BEPS project range from a strong commitment by participating countries to the weak form of commitment. These different actions exhibit different levels of complementarity across countries. In particular, actions such as countering harmful tax practice (Action 5), prevention of tax treaty abuses (Action 6), country-by-country reporting (Action 13), and mutual agreement procedure (Action 14) are called "minimum standards," and they display the different level of complementarity than other actions, such as the limitation on interest deduction (Action 4) and transfer pricing limitation (Actions 8–10). 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 of the public good provision analyzed by Hirshleifer (1983, p. 373). We formalize the collective-action problems of enforcement by using Hirshleifer's (1983) social composition function. Depending on the forms of enforcement contributions, different aggregation technology will prevail, varying from perfect substitutability to perfect complementarity (the weakest-link). We show that stronger complementarity in enforcement, together with the asymmetry among countries, induces a race to the bottom in enforcement (Lemma 3), leading to undertaxation for both countries. As a result, they are both willing to cooperate on enforcement to raise their tax revenues (Proposition 1).

The second remedy for the lack of cooperation is tax leadership, which has attracted research interest since Kempf and Rota-Graziosi (2010).[4] Following Hindriks and Nishimura (2015, 2017), we consider the case in which the high-tax country (the large country) leads.[5] Compared with simultaneous tax choice, a widening of the tax gap and a reduction of the tax-revenue gap occur (Lemma 4). The latter, which is a new feature of the tax leadership to our knowledge, reduces the conflict of interest on enforcement levels and hence increases the benefit of enforcement cooperation (Proposition 2). Our analysis also shows that enforcement cooperation is impossible when asymmetry is too high and notably in the presence of tax havens.

A typical debate on corporate income taxes involves the choice between separate accounting (SA) and formula apportionment (FA). However, a shift from SA to FA does not solve fiscal spillover problems and it may even aggravate them (Nielsen et al., 2010). More fundamentally, each country has preferred sharing rules (e.g., sales, capital, and labor) so that a group of countries would never agree on allocation rules. Notably, such a disagreement appears in the current debate on taxation of digital services, in setting permanent establishment (PE) concept

[4] 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 reaction to the United States 1986 Tax Reform Act.

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

on digital products. Another route when the scope for enforcement cooperation is limited is to consider a shift from source-based to residence-based taxation. However, 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.

The main focus of the previous studies was to understand how transfer prices are affected by international tax differences and tax systems (see, e.g., Amerighi & Peralta, 2010; Devereux et al., 2008; Huizinga & Laeven, 2008; Kind et al., 2005; Klassen & Laplante, 2012; Nielsen et al., 2008; Swenson, 2001). In addition, some evidence showed that transfer pricing regulations significantly mitigate profit shifting (e.g., Bartelsman & Beetsma, 2003). As in the current paper, several studies modeled tax competition augmented with enforcement choices.[6] 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, given that only the latter 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.[7] Stöwhase (2013) showed that permitting profit shifting may soften the tax competition for attracting capital, since the profit can be relocated ex post to the low-tax country.[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. Keen and Slemrod (2017) developed a framework to derive optimal tax enforcement. They provided a guideline for optimal compliance gap (which is nonzero) based on the enforcement elasticity of the tax revenue. In this paper, we consider an international setting where countries compete in tax rates and enforcement levels. Using a spatial econometric approach, Durán-Cabré et al. (2015) provided evidence of strategic complementarities between regional administrations with respect to audit policies among Spanish regional governments.

The rest of the paper is organized as follows. Section 2 describes the model, firms' profit shifting and the governments' tax choices. Section 3 compares the noncooperative and cooperative enforcement regimes, and examines the benefit of cooperation in relation to enforcement complementarity. Section 4 examines tax leadership. Section 5 concludes with some policy implications. 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 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:

[6]There is also related literature on tax competition with amenities and public infrastructure (see Dhillon et al., 2007; Hindriks et al., 2008), where the main interest is on the interplay between public infrastructure/amenities provision and tax competition.

[7]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.

[8]The argument is similar to those by Hong and Smart (2010) and Johannesen (2010), in that the presence of tax havens may soften the tax competition. However, such theoretical possibility is at odds with empirical evidence on tax planning and policy discussion on the limited scope for local taxation of global organizations in the presence of profit shifting.

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

That is, country 1 is the *large* country because (i) α_i, which represents the difference in countries' income, is higher, or (ii) the supply cost c_i is lower.

From the production decisions $\left(q_i^a, q_i^b\right)$, in country $i = 1, 2$, firm $k = a, b$ generates $\pi_i^k = \left\{p_i\left(q_i^a + q_i^b\right) - c_i\right\}q_i^k$ 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 $\left(\tilde{\pi}_1^k + \tilde{\pi}_2^k = \pi_1^k + \pi_2^k\right)$. 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 and nonfiscally deductible concealment cost $C(\pi_i^k, \tilde{\pi}_i^k)$, which is widely used in the literature:[9]

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

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.[10] 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 constant elasticity of substitution (CES) formula:

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

The enforcement technology (2) is exogenous. 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.

The tax revenue in country i is:

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

[9] 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). See also Huizinga and Laeven (2008) and Amerighi and Peralta (2010), for a slightly different specification.

[10] 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.

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 $\left(c(e_i) = \eta\frac{(e_i)^2}{2}\right)$ 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.[11]

2.2 | Profit shifting by firms

We consider a three-stage game with the following sequence of events. In the first stage, both countries set their enforcement efforts. In the second stage, both countries choose their tax rates. In the third stage, MNEs 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.

Regarding enforcement and tax timing, we assume that the enforcement efforts are chosen first and taxes are chosen later.[12] 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 more easily (see Bacchetta & Espinosa, 1995; Bucovetsky & Haufler, 2008; Peralta et al., 2006, and Keen & Konrad, 2013). 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.

The model is solved by backward induction. In this section, 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, $\left(q_1^k, q_2^k\right)$ and the profit to report, $\left(\tilde{\pi}_1^k, \tilde{\pi}_2^k\right)$, 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 $\tilde{\pi}_1^k + \tilde{\pi}_2^k = \pi_1^k + \pi_2^k$.

[11] In most of the paper, we can restrict $\eta = 1$ without loss of generality. The exception is in Section 4 when the high-tax country takes the leadership, because then taxes are higher. A sufficient condition for interior tax equilibrium is $\eta \geq 3$ in that specific case.

[12] In our discussion paper (Hindriks & Nishimura, 2018), we considered the scenario that the tax rate and enforcement effort are chosen simultaneously and noncooperatively. We showed that there is no equilibrium with positive taxes in pure strategy. In Section 3.2, we briefly discuss the case where taxes are chosen first and the enforcement level is chosen later.

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

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

$\left(\text{or } \frac{\gamma_1^2}{\gamma_1^2 + \gamma_2^2} = \frac{1+\epsilon}{2} \in \left[1/2, 1\right)\right)$. We normalize $\beta = \frac{2}{9}\left(\gamma_1^2 + \gamma_2^2\right)$ so that the total profit $\sum_{k=a,b}\left(\pi_1^k + \pi_2^k\right)$ in Lemma 1 below is normalized to 1. Profit taxes do not change supply and aggregate profit. However, profit taxes change the distribution of profit shares (reported profit) between countries via profit shifting. Let $\tilde{\pi}_i = \tilde{\pi}_i^a + \tilde{\pi}_i^b$ be the total reported profit in country i. In the Appendix we show the following:

$$\begin{gathered} \pi_1^k = \frac{1+\epsilon}{4}, \quad \pi_2^k = \frac{1-\epsilon}{4}(k = a, b), \\ \tilde{\pi}_1^a + \tilde{\pi}_1^b = \frac{1+\epsilon}{2} - \frac{t_1 - t_2}{2\delta(e)} \equiv \tilde{\pi}_1(t, e), \quad \tilde{\pi}_2^a + \tilde{\pi}_2^b = \frac{1-\epsilon}{2} + \frac{t_1 - t_2}{2\delta(e)} \equiv \tilde{\pi}_2(t, e). \end{gathered} \tag{4}$$

The implications of (4) are as follows:

Lemma 1. *The reported profit in country i consists of (a) actual profits $\pi_i = \pi_i^a + \pi_i^b$ that depend in the market size ϵ and (b) the amount of profit shifting $\pi_i - \tilde{\pi}_i$. The latter is proportional to the tax difference $t_i - t_j$ and inversely proportional to the total enforcement $\delta(e)$.*

Proof. See the Appendix. □

Lemma 1 implies that for any given enforcement level, the large country has a larger tax base (actual profit) but the small country can attract part of the tax base by taxing less. The implication to our model is that the large country perceives its tax base to be less elastic than the small country.[13] This is a common feature of asymmetric tax competition models which as we now show will induce the small country to tax less in equilibrium.

From Lemma 1, given the equilibrium profit shifting, country i's tax revenue net of the enforcement cost is:

$$W_i = t_i \tilde{\pi}_i(t_i, t_j, e) - \eta \frac{(e_i)^2}{2} = t_i\left(\frac{1+\epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)}\right) - \eta\frac{(e_i)^2}{2}, \tag{5}$$

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

2.3 | Tax choices

In the second stage of the game, each country noncooperatively chooses its own tax rate $t_i(i = 1, 2)$to maximize (5). The first-order conditions are:

$$\frac{\partial W_i}{\partial t_i} = \frac{1+\epsilon_i}{2} - \frac{t_i - t_j}{2\delta(e)} + t_i\frac{-1}{2\delta(e)} = 0. \tag{6}$$

[13] The tax base elasticities at $t_1 = t_2 = t$ are $-\frac{\partial\tilde{\pi}_1}{\partial t_1}\frac{t_1}{\tilde{\pi}_1} = \frac{t}{\delta(e)(1+\epsilon)}$ and $-\frac{\partial\tilde{\pi}_2}{\partial t_2}\frac{t_2}{\tilde{\pi}_2} = \frac{t}{\delta(e)(1-\epsilon)}$.

The second-order conditions are satisfied. They yield the following equilibrium taxes, denoted by $\left(t_1^N(e), t_2^N(e)\right)$:

$$t_1^N(e) = \delta(e)\left(\frac{3+\epsilon}{3}\right) \quad \text{and} \quad t_2^N(e) = \delta(e)\left(\frac{3-\epsilon}{3}\right). \tag{7}$$

Reflecting its smaller tax base elasticity, the large country sets higher taxes. The tax rate differential $t_1^N(e) - t_2^N(e) = 2\frac{\delta(e)\epsilon}{3} > 0$ widens with greater asymmetry. What is more striking is that the tax gap is also proportional to the enforcement level. From Lemma 1, for $\tilde{\pi}_i^N \equiv \tilde{\pi}_i\left(t_1^N(e), t_2^N(e), e\right)$, we have:

$$\tilde{\pi}_1^N = \frac{1+\epsilon}{2} - \frac{t_1^N(e) - t_2^N(e)}{2\delta(e)} = \frac{1}{2}\frac{3+\epsilon}{3} \quad \text{and} \quad \tilde{\pi}_2^N = \frac{1}{2}\frac{3-\epsilon}{3}, \quad \text{so} \quad \pi_1 - \tilde{\pi}_1^N = \frac{\epsilon}{3}. \tag{8}$$

That is, in equilibrium, the profit shares (reported profits) as well as the profit shifting $\pi_1 - \tilde{\pi}_1^N$ are independent of enforcement levels (e_1, e_2) and enforcement technology $\delta(e)$. The higher enforcement simply scales up the tax gap by the same proportion. They cancel each other so that the extent of profit shifting in equilibrium is unchanged. However, greater market asymmetry increases the profit share of the high-tax country. The tax revenues $R_i^N(e) = t_i^N(e)\tilde{\pi}_i^N (i = 1, 2)$ are as follows:

$$R_1^N(e) = \frac{\delta(e)}{2}\left(\frac{3+\epsilon}{3}\right)^2 \geq R_2^N(e) = \frac{\delta(e)}{2}\left(\frac{3-\epsilon}{3}\right)^2 \quad \text{for all} \quad \epsilon \geq 0. \tag{9}$$

The total reported profit $\tilde{\pi}_1^N + \tilde{\pi}_2^N = 1$ is constant. The total revenue $\sum_i R_i^N(e) = \sum_i t_i^N(e)\tilde{\pi}_i^N$ is a weighted average of the tax rates with weights given by the profit shares. Greater asymmetry increases both the profit share of the high-tax country $\tilde{\pi}_1^N$ and its tax rate $t_1^N(e)$. As a result, given e, the total revenue increases with asymmetry.

3 | ENFORCEMENT CHOICES AND COMPLEMENTARITY

3.1 | Simple illustrations

In the first stage, the governments in each country choose their enforcement effort levels, taking into account the behavior in the subsequent stages. We first examine noncooperative enforcement choices, where each country chooses $e_i (i = 1, 2)$ simultaneously and independently. Let $\left(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon)\right)$ be the enforcement level at the noncooperative equilibrium with ρ and ϵ given, and let $\delta^N(\rho, \epsilon) \equiv \delta\left(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon)\right)$ be the corresponding overall enforcement level.

3.1.1 | Illustration 1: Market asymmetry

For an illustration, we first consider the case of perfect substitutability between enforcement choices ($\rho = -1$) so that the overall enforcement is the arithmetic mean $\delta_{-1} \equiv 0.5e_1 + 0.5e_2$. We also set $\eta = 1$ in this section for simplicity. Given e_j, country i maximizes $W_i^N(e_i, e_j) = R_i^N(e_i, e_j) - \frac{(e_i)^2}{2}$, where $R_i^N(e)$'s are given in (9). The first-order condition with respect to country i's enforcement choice is given by $\left(\frac{3+\epsilon_i}{6}\right)^2 - e_i = 0$ with $\epsilon_1 = \epsilon = -\epsilon_2$. Hence, under perfect enforcement substitutability, country i's enforcement choice is independent of $e_j (j \neq i)$. The equilibrium is given by:

$$e_1^N(-1, \epsilon) = \left(\frac{3+\epsilon}{6}\right)^2, \quad e_2^N(-1, \epsilon) = \left(\frac{3-\epsilon}{6}\right)^2, \quad \delta^N(-1, \epsilon) = \frac{9+\epsilon^2}{36}. \tag{10}$$

Greater market asymmetry shifts both profit shares and enforcement levels from the low-tax country to the high-tax country. Since the benefit of enforcement is proportional to the tax revenue which is convex in the degree of asymmetry, increasing asymmetry induces the high-tax country to raise its enforcement to a greater extent than the reduction by the low-tax country. Therefore, overall enforcement (which is the average of the enforcement efforts when $\rho = -1$) increases with asymmetry.

Plugging (10) into (9), we find that the total revenue $R_1^N(e^N(-1, \epsilon)) + R_2^N(e^N(-1, \epsilon))$ is increasing with asymmetry, taking the value 1/4 when countries are symmetric ($\epsilon = 0$), and $(1/4 + 1/36)\cdot 10/9$ when $\epsilon = 1$. Here, the high-tax country can compete with the tax haven not only with the tax rates but also with greater enforcement levels. As a result, total revenue is higher in the presence of a tax haven.

As we now show in the next illustration, the enforcement competition changes qualitatively when the enforcement efforts are complements.

3.1.2 | Illustration 2: Enforcement complementarity

Next, we consider enforcement complementarity given by the following technology ($\rho = 0$): $\delta_0 \equiv e_1^{0.5} e_2^{0.5}$, where the overall enforcement takes the form of geometric mean. We set $\eta = 1$ in this section, too.

The first-order condition with respect to e_i yields country i's enforcement reaction function given by $r_i(e_j) = e_j^{\frac{1}{3}} \left(\frac{3+\epsilon_i}{6}\right)^{\frac{4}{3}}$. For $\rho > -1$, the enforcement reaction functions are upward-sloping, that is, there is strategic complementarity in enforcement efforts (see Figure 1).

For $\epsilon = 0$, the equilibrium enforcement level (the intersection of the enforcement reaction functions) is $e_i^N(-1, 0) = 1/4 = e_i^N(0, 0) \equiv e^*$ for $i = 1, 2$, and the overall enforcement is $\delta_0(e^N(-1, 0)) = 1/4 = \delta_{-1}(e^N(0, 0))$. Thus, enforcement complementarity has no impact when countries are identical.

The matters are different when countries are asymmetric. For $\epsilon > 0$, the large country 1 will set higher tax than country 2, inducing profit shifting to country 2. As a result, country 1 will have greater incentive for enforcement than country 2. Enforcement levels are no longer aligned, and the symmetry between the two enforcement technologies is lost. At the intersection of the enforcement reaction function, we have $\delta^N(0, \epsilon) = \frac{(3+\epsilon)(3-\epsilon)}{36}$, which is smaller than $\delta^N(-1, \epsilon)$ in (10). To see the difference between the two technologies, consider a mean preserving dispersion of the enforcement choices, from $\mathbf{E} = (e^*, e^*)$ to $\mathbf{C} = (e^* + \Delta, e^* - \Delta)$ in Figure 1. When enforcement efforts are perfect substitutes, the overall enforcement is unchanged ($\delta_{-1}(e^* + \Delta, e^* - \Delta) = \delta_{-1}(e^*, e^*)$). When enforcement efforts are complements, the overall enforcement is reduced: $\delta_0(e^* + \Delta, e^* - \Delta) < \delta_0(e^*, e^*)$. Therefore, complementarity reduces the effectiveness of enforcement dispersion. We show in the next section that, when $\epsilon > 0$, the total enforcement level decreases with greater complementarity.

In contrast with (10) for $\rho = -1$, the equilibrium overall enforcement is decreasing in ϵ for the case of $\rho = 0$, since the impact of $e_2^N(0, \epsilon)$'s decrease with respect to asymmetry[14] has the greater weight in the geometric mean than in the case of perfect substitution. Plugging $\delta^N(0, \epsilon)$ into $\delta(e)$ of (9), the equilibrium total revenue $R_1^N(e^N(0, \epsilon)) + R_2^N(e^N(0, \epsilon))$ is decreasing in ϵ:[15] the effect of $\delta^N(0, \epsilon)$ decreasing in ϵ is dominant in the total revenue, so that asymmetry is not beneficial for the total revenue.

We showed three features for $\rho = 0$: (a) for $\epsilon = 0$, enforcement complementarity does not have an impact on the equilibrium overall enforcement, (b) complementarity makes dispersed enforcement less effective ($\delta^N(\rho, \epsilon)$ decreases in ρ for $\epsilon > 0$), and (c) $\delta^N(\rho, \epsilon)$ decreases in ϵ. These features continue to hold for sufficiently high ρ, including the polar case of perfect complementarity $\rho = \infty$.

3.2 | Noncooperative enforcement choices: General cases

Now, we move on to general cases. First, we show that $e_1^N(\rho, \epsilon) > e_2^N(\rho, \epsilon) > 0$ for all $\rho < \infty$ and $\epsilon > 0$.

Lemma 2. *The high-tax country also sets a higher level of enforcement.*

Proof. See the Appendix. □

Lemma 2 illustrates the synergy between market size and enforcement. 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.[16]

In the Appendix, we show the following:

Lemma 3.

(i) Under asymmetry, more enforcement complementarity reduces the dispersion of enforcement by shifting enforcement from the high-tax country to the low-tax country. The overall enforcement is decreasing with complementarity.

(ii) More asymmetry widens the enforcement gap. There exists a threshold level of complementarity ($\tilde{\rho} > -1$) such that overall enforcement decreases with asymmetry if and only if $\rho > \tilde{\rho}$.

---

[14] $e_2^N(0, \epsilon) = \frac{(3-\epsilon)^{3/2}(3+\epsilon)^{1/2}}{6^2}$. In our discussion paper (Hindriks & Nishimura, 2018), we showed that $e_2^N(\rho, \epsilon)$ decreases in ϵ for all ρ and ϵ.

[15] $R_1^N(e^N(0, 0)) + R_2^N(e^N(0, 0)) = 1/4\left(= R_1^N(e^N(-1, 0)) + R_2^N(e^N(-1, 0))\right)$ when countries are symmetric, and $R_1^N(e^N(0, 1)) + R_2^N(e^N(0, 1)) = \frac{2}{9}\frac{10}{9}$ when $\epsilon = 1$.

[16] 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 ϵ is sufficiently high, and (ii) both the equilibrium taxes and enforcement efforts are lower than those obtained in (7) and Lemma 2 (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 ($e_1 > 0$ 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 (7).

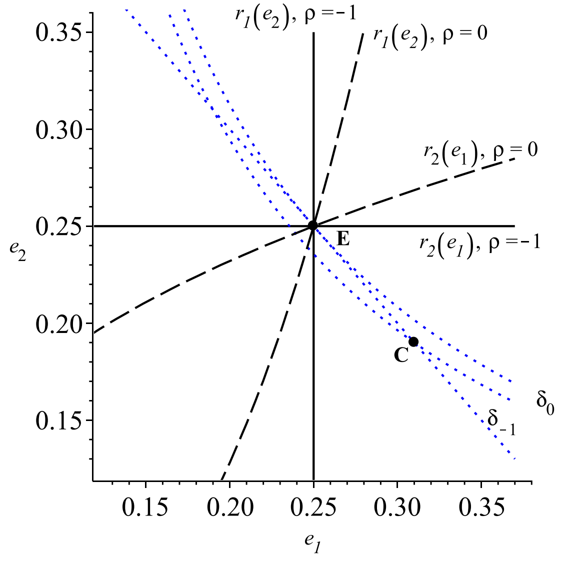

FIGURE 1 Enforcement reaction functions for $\rho = -1, \epsilon = 0$ (solid curves) and for $\rho = 0, \epsilon = 0$ (dashed curves)

Proof. See the Appendix. □

Figure 2 illustrates the enforcement reaction function of country i for the cases of $\rho = -1, \epsilon = 0.25$ and $\rho = 0, \epsilon = 0.25$. The dotted curve **PW** illustrates the locus of the equilibrium enforcement efforts (the intersection of the reaction functions) for $\epsilon = 0.25$. From $\mathbf{P} = \left(e_1^N(-1, \epsilon), e_2^N(-1, \epsilon)\right)$, greater complementarity (larger ρ) reduces the best-response

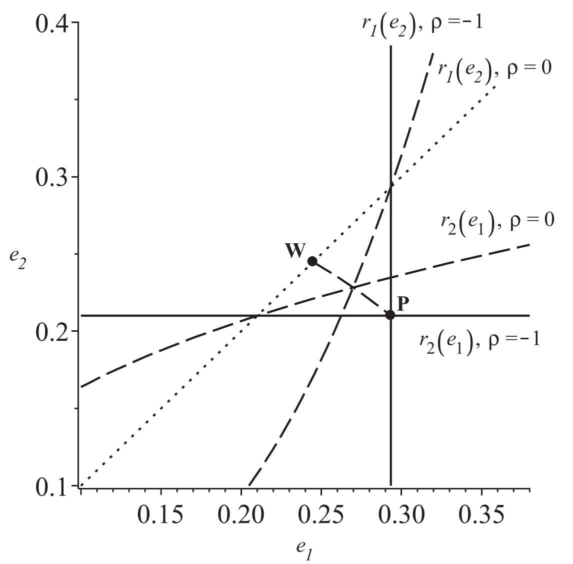

FIGURE 2 Enforcement reaction functions for $\rho = -1, \epsilon = 0.25$ (solid curves) and for $\rho = 0, \epsilon = 0.25$ (dashed curves). The curve **PW** is the locus of $\left(e_1^N(\rho, \epsilon), e_2^N(\rho, \epsilon)\right)$ for $\epsilon = 0.25$ and $\rho \in [-1, \infty)$

effort of country 1 and increases that of country 2. The reason is that for $e_1 > e_2$, greater complementarity lowers the marginal productivity of e_1 and increases the marginal productivity of e_2. As discussed before, complementarity increases the efficiency loss of enforcement dispersion and so it induces countries to bring their enforcement levels closer to each other. However, this (partial) alignment is not sufficient to offset the efficiency loss from enforcement dispersion, so that, as an extension of $\delta^N(0, \epsilon) < \delta^N(-1, \epsilon)$ in the previous section, overall enforcement declines with complementarity.

Greater asymmetry shifts outwards the best response of country 1 and shifts downwards the best response of country 2, so it increases the gap in enforcement levels. The overall enforcement $\delta^N(\rho, \epsilon)$ is increasing in ϵ when ρ is sufficiently low (as in the perfect-substitution case), reflecting the fact that equilibrium enforcement is proportional to the country's tax revenue which is convex in ϵ. However, when ρ is sufficiently high, the efficiency loss from the increased dispersion of enforcement becomes dominant, and the overall enforcement ($\delta^N(\rho, \epsilon)$) decreases with asymmetry.

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 keeping with the current OECD framework to reinforce enforcement cooperation in which each country still can freely choose tax rates, we assume that countries set taxes (t_i) noncooperatively. Therefore, countries choose $e = (e_1, e_2)$, anticipating the noncooperative tax game $\left(t_1^N(e), t_2^N(e)\right)$ and tax revenues $\left(R_1^N(e), R_2^N(e)\right)$ in (9). That is:

$$\max_{e_1, e_2} \sum_i \left(R_i^N(e_i, e_j) - \eta \frac{(e_i)^2}{2}\right). \tag{11}$$

As the solution of (11), when $\eta = 1$, we obtain $\hat{e}_1^N(\epsilon) = \hat{e}_2^N(\epsilon) = \frac{9+\epsilon^2}{18} = \hat{\delta}^N(\epsilon)$, 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, and with the CES technology, the enforcement effectiveness becomes independent of ρ. When $\rho = -1$, compared with the noncooperative solution (10), 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. For $\rho > -1$ and $\epsilon > 0$, the efficiency loss from enforcement dispersion reinforces the difference between cooperative and noncooperative total enforcement.

Let $W_i^N(e^N(\rho, \epsilon))$ and $W_i^N(\hat{e}^N(\epsilon))(i = 1, 2)$ be the welfare levels in the noncooperative and cooperative regimes, respectively. Decomposing the benefit and the cost of cooperation for country i gives:

$$W_i^N(\hat{e}^N(\epsilon)) - W_i^N(\rho, \epsilon) = \underbrace{(\hat{\delta}^N(\epsilon) - \delta^N(\rho, \epsilon)) \frac{(3+\epsilon_i)^2}{18}}_{\text{benefit (increased revenue)}} - \underbrace{\frac{\eta}{2}((\hat{e}_i^N(\epsilon))^2 - (e_i^N(\rho, \epsilon))^2)}_{\text{cost (increased effort)}}.$$

Under symmetry, both countries unambiguously benefit from enforcement cooperation since the aggregate welfare is maximized and the surplus is equally shared. Under asymmetry, the surplus is no longer equally shared. Specifically, cooperation involves enforcement harmonization which requires more enforcement effort from the low-tax country than its level under noncooperation ($\hat{e}_2^N(\epsilon) > e_2^N(\rho, \epsilon)$). When asymmetry is large, the cost of this extra enforcement exceeds the benefit, and the low-tax country prefers not to cooperate. This feature corresponds to a small country being unwilling to cooperate when tax rates are the only instrument of tax competition (Bucovetsky, 1991; Wilson, 1991). However, we go further than this.

Under asymmetry, higher complementarity lowers the overall enforcement (Lemma 3) which exacerbates tax competition and reduces equilibrium taxes for both countries. In addition, country 2 increases noncooperative enforcement effort with greater complementarity, due to the enforcement-matching effect shown in Lemma 3 and Figure 2. Therefore, the low-tax country is more likely to agree on a cooperative enforcement level that enables better inspection and information sharing. In contrast, when enforcement efforts are substitutes, the low-tax country free rides on the enforcement of the high-tax country whereas the dispersion of enforcement levels has less impact on the overall enforcement level. Higher level of overall enforcement attenuates the extent of the tax competition, and taxes at noncooperative equilibrium increase. When the relevant parameter to be coordinated is tax enforcement, the following proposition states the conditions under which enforcement cooperation is mutually beneficial.

Proposition 1. *Under symmetry, both countries benefit from enforcement cooperation. Under asymmetry, the high-tax country unambiguously gains from enforcement cooperation. As for the low-tax country, there exists a cutoff level of asymmetry below which it benefits from cooperation. This cutoff level is below 1 (excluding the case of tax haven), and it is increasing with complementarity of the enforcement technology.*

Proof. See the Appendix. □

For a given level of enforcement complementarity (ρ), country 2 prefers the noncooperative regime when asymmetry is above a certain cutoff value ($W_2^N(\hat{e}^N(\epsilon)) < W_2^N(e^N(\rho, \epsilon))$ iff $\epsilon > \epsilon^N(\rho)$). This cutoff level $\epsilon^N(\rho)$ is increasing in ρ. Lower overall enforcement ($\delta(e)$) seems to benefit the low-tax country, but it induces the high-tax country to match its tax to the rival so as to scale down the equilibrium tax rates in (7). As a result, the extent of equilibrium profit shifting in (8) is invariant with respect to enforcement levels in our model. Therefore, lower enforcement does not translate into more profit shifting but rather into lower equilibrium taxes, making both countries worse off. In the (ρ, ϵ) space, Figure 3 indicates where cooperation is mutually beneficial (below/right of the curve $\epsilon^N(\rho)$).

Proposition 1 suggests some policy implications. The low-tax country is more likely to accept enforcement cooperation (harmonization) when enforcement efforts are complements and the asymmetry in market size is not high. This would be the case if, for instance, the provision of tax-relevant information by the host (low-tax) country is crucial for the effectiveness of the overall enforcement. In contrast, if the enforcement actions are substitute with respect to the transfer pricing documentation from each country or the number of tax officials in each country, then the low-tax country may prefer the free-rider benefit and forego the benefit of enforcement cooperation.

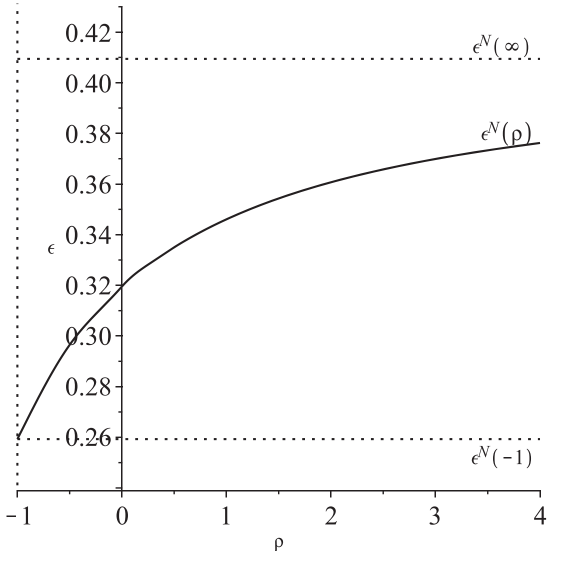

FIGURE 3 Critical value of asymmetry $\epsilon^N(\rho)$

4 | TAX COMMITMENT

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 (hereafter referred to as the high-tax leadership) follows Hindriks and Nishimura (2015, 2017).[17]

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. Along country 2's tax reaction function $t_2 = t_2^r(t_1; e)$, country 1, as the Stackelberg leader, chooses t_1. This interaction yields the equilibrium tax rates denoted by $\left(t_1^S(e), t_2^S(e)\right)$ 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 4. *For any given level of enforcement $e = (e_1, e_2)$:*

(i) High-tax leadership induces wider tax gaps and higher tax revenues relative to the Nash tax competition.

(ii) High-tax leadership reduces the revenue gap relative to the Nash tax competition, when asymmetry is sufficiently large.

(i) High-tax leadership induces wider tax gaps and higher tax revenues relative to the Nash tax competition.

[17] At the end of this section, we briefly discuss the case of the low-tax leadership.

(ii) High-tax leadership reduces the revenue gap relative to the Nash tax competition, when asymmetry is sufficiently large.

Proof. See the Appendix. □

Under tax leadership, the tax leader selects its most-preferred tax rate along the best-response of the tax follower. Given that taxes are strategic complements as in the conventional models, equilibrium tax rates and revenues are higher than under the Nash tax competition, and the follower benefits more from this tax commitment (Hindriks & Nishimura, 2015, 2017). Hence, if the high-tax country takes the lead, the revenue of the low-tax country will increase more than that of the tax leader.[18]

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

Lemma 5. *High-tax leadership increases enforcement efforts and reduces their dispersion relative to the Nash tax competition, when asymmetry is sufficiently large.*

Proof. See the Appendix. □

Next, we derive the cooperative solution. 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 by choosing $\left(t_1^S(e), t_2^S(e)\right)$. 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 $e = \left(\hat{e}_1^S, \hat{e}_2^S\right)$. Under the Nash game (with $\rho = -1$) in Proposition 1 (Figure 3), the low-tax country prefers the noncooperative regime $(W_2^N(\hat{e}^N(\epsilon)) < W_2^N(e^N(-1, \epsilon)))$ if and only if $\epsilon > \epsilon^N(-1)$. In comparison to that case, we show the following:

Proposition 2. *High-tax leadership increases the chances of cooperation on enforcement: the cutoff level of asymmetry below which the low-tax country gains is higher than that under the Nash tax competition. This cutoff level is below 1 (excluding the case of tax haven).*

Proof. See the Appendix. □

Lemma 4 indicates that tax leadership by the high-tax country will act as a fiscal equalizer that counteracts the size asymmetry. As a result, countries are more likely to cooperate on

[18] Since the slopes of the tax response functions are less than one ($\partial t_2^r(t_1; e)/\partial t_1 < 1$), the tax increase by the tax follower is less than that of the tax leader. This widening of the tax gap induces more profit shifting, which benefits the low-tax country. In the Appendix, we show that $0 < R_1^S(e) - R_1^N(e) < R_2^S(e) - R_2^N(e)$.

enforcement. To see why this happens, decomposing the benefit and the cost of cooperation for country i gives:

$$W_i^S(\hat{e}^S) - W_i^S(e^S) = (R_i^S(\hat{e}^S) - R_i^S(e^S)) - \frac{\eta}{2}((\hat{e}_i^S)^2 - (e_i^S)^2).$$

Note that cooperation produces a surplus and so at least one country must derive a benefit. Since the revenue gap and enforcement gap are reduced under tax leadership by the high-tax country, as shown in Lemmas 4 and 5, the surplus of cooperation is more equally shared. Therefore, cooperation is more likely to be agreed unanimously under tax leadership. This result continues to hold for all ρ in (2).[19] The policy implication is that high-tax leadership is a more acceptable form of coordination for the low-tax country than tax harmonization, which eliminates the benefit of profit shifting.

Note that $\epsilon = 1$ (no production taking place in country 2), which can be interpreted as country 2 being an offshore tax haven, is excluded from the case of beneficial cooperation in Propositions 1 and 2. Consistent with the real world situation, it is difficult to persuade tax-haven countries to cooperation.

Finally, we discuss an alternative timing of low-tax leadership.[20] The result of Proposition 2 is reversed when the low-tax country takes the lead. The reason is as follows. When the low-tax country is the tax leader, the tax-follower benefit is now shifted to the high-tax country. Leadership by the low-tax country widens the fiscal gap, exacerbating the initial asymmetry: for the equilibrium enforcement effort $e_i^M (i = 1, 2)$ under country 2's leadership, we have $e_1^M - e_2^M > e_1^N(-1, \epsilon) - e_2^N(-1, \epsilon)$ for all ϵ. The cooperation threshold level of asymmetry is now lower, making enforcement cooperation less likely under the low-tax leadership.[21] This last result has an immediate policy implication. Suppose that we impose minimum taxation forcing the low-tax country to raise its tax rate. It may have similar effects on enforcement as low-tax leadership. The reason is as follows. This minimum tax would act as a commitment of the low-tax country to tax more that triggers higher taxes from the high-tax country. Since the high-tax country gets the benefits corresponding to the tax-follower benefit, the low-tax country would be less likely to cooperate on enforcement. Hence, tax harmonization from below, including the minimal tax standards, would reduce the scope for enforcement cooperation. This point is similar to the argument in the tax-competition literature that tax harmonization could induce countries to adopt lenient enforcement policies.

5 | CONCLUSION AND POLICY IMPLICATIONS

This paper discussed the scope for enforcement cooperation in the context of un-coordinated tax choices among asymmetric countries. Under the profit shifting behavior by MNEs seeking tax minimization through legal exploitation of tax gaps across countries, enforcement cooperation is represented as a simple strategic game between

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

[20] 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).

[21] The cutoff value of asymmetry sustaining cooperation is even smaller than that in Proposition 1. That is, for the cooperative effort level $\left(\hat{e}_1^M, \hat{e}_2^M\right)$, we have $W_2^M\left(\hat{e}_1^M, \hat{e}_2^M\right) < W_2^M\left(e_1^M, e_2^M\right)$ for all $\epsilon > \epsilon^N(-1)$. Moreover, we can show that the high-tax leadership equilibrium Pareto dominates the low-tax leadership equilibrium under sufficient asymmetry: $W_2^M\left(e_1^M, e_2^M\right) < W_2^S\left(e_1^S, e_2^S\right)$ for all ϵ, and $W_1^M\left(e_1^M, e_2^M\right) < W_1^S\left(e_1^S, e_2^S\right)$ if ϵ is sufficiently high.

revenue-maximizing governments. We designed a model that captures the multilevel strategic interactions and provides closed-form solutions to assess the impact of various enforcement technologies and tax commitments on the benefits and costs of enforcement cooperation. Our first main finding is that enforcement complementarity facilitates cooperation in the sense that the low-tax country is more willing to agree. This result is somewhat striking because stronger complementarity gives more power to the low-tax country in setting a noncooperative enforcement level. There are two reasons why the low-tax country may prefer enforcement cooperation. First, lower enforcement does not translate into more profit shifting but rather into lower equilibrium tax revenues. Second, cooperation leads to enforcement harmonization, making the overall enforcement more effective compared to dispersed enforcements. A key feature of enforcement complementarity is that the efficiency loss from enforcement dispersion at noncooperative equilibrium is higher.

The policy implication from this result is the following. To facilitate the participation of the low-tax country in the enforcement agreement, we should target the enforcement actions that are complements rather than substitutes. Such actions can be identified as those on which the low-tax country has more influence. For instance, the low-tax country may be able to exercise some veto power towards tighter enforcement on the automatic exchange of information during a mutual agreement procedure on the taxable incomes of the MNEs. A simple yet powerful strategy to assess the complementarity of a given action is to estimate the efficiency loss (in terms of the revenue) from dispersed actions among countries. For instance, OECD's (2015) Action 11 (measuring and monitoring) involves collecting and analyzing fiscal data, which is something each country can do to a different extent without hampering the overall effect. This is also true for actions that involve strengthened guidelines for transfer pricing (Actions 8–10): it is not required for all countries to impose the same strict guidelines for this action to be effective. In contrast, the actions involving “minimum standards” require similar actions from the different countries to be effective. Also, Action 1 on digitalization requires similar actions among countries, because this action requires that no special regime should be created for digital taxation.

Our second main result is related to tax commitment. We showed that tax leadership by the high-tax country (high-tax leadership) facilitates enforcement cooperation by the low-tax country. The reason is that such tax leadership acts as a fiscal equalizer that mitigates the conflict of interest on enforcement between the low-tax and high-tax countries. Interestingly, this type of coordination strategy is different from tax harmonization, which eliminates the benefit of profit shifting to the small country. Conversely, under tax leadership by the low-tax country (low-tax leadership), the benefit of the tax follower is shifted to the high-tax country which increases the extent of the disagreement on enforcement. The policy implication is that a tax commitment from the low-tax country (such as minimal taxation) will make that country less willing to agree on enforcement. Lastly, our results suggest that it is impossible to get a tax haven to agree on enforcement cooperation.

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.

ACKNOWLEDGMENTS

We would like to thank the Editor (Helmuth Cremer), the Associate Editor, and the referees for their useful comments. This paper was presented at the International Symposium of Urban Economics and Public Economics (Osaka University), the 3rd Belgian–Japanese Public Finance Workshop (CORE), and seminars in École normale supérieure Paris-Saclay and University of Tokyo. The authors wish to thank the participants, particularly David Agrawal, Pierre Boyer, Nicolas Gravel, Andreas Haufler, Hubert Kempf, Mutsumi Matsumoto, Susana Peralta, and Emmanuelle Taugourdeau, for their useful comments. Financial support from the Obayashi Foundation and the Japan Legislative Society Foundation is gratefully acknowledged. In addition, Yukihiro Nishimura is grateful for financial support from the Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (JP15K03511, JP18H00866). The authors also thank Edanz Group (https://www.edanzediting.com/ac) for editing a draft of this manuscript.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Jean Hindriks http://orcid.org/0000-0001-7302-7122
Yukihiro Nishimura http://orcid.org/0000-0001-9239-1819

REFERENCES

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

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

Bacchetta, P., & Espinosa, M. P. (1995). Information sharing and tax competition among governments. *Journal of International Economics*, *39*, 103–121.

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

Becker, J., Johannesen, N., & Riedel, N. (2020). Taxation and the allocation of risk inside the multinational firm. *Journal of Public Economics*, *183*, 104138.

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

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.

Clausing, K. (2003). Tax-motivated transfer pricing and US intrafirm trade prices. *Journal of Public Economics*, *87*, 2207–2223.

Clausing, K. (2009). Multinational firm tax avoidance and tax policy. *National Tax Journal*, *62*, 703–725.

Cremer, H., & Gahvari, F. (2000). Tax evasion, fiscal competition and economic integration. *European Economic Review*, *44*, 1633–1657.

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

Dhillon, A., Wooders, M., & Zissimos, B. (2007). Tax competition reconsidered. *Journal of Public Economic Theory*, *9*, 391–423.

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.

European Commission. (2020). *The Fiscalis 2020 programme*. Taxation and Customs Union, European Commission, Brussels. https://ec.europa.eu/taxation_customs/fiscalis-programme_en

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

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.

Haufler, A., & Schjelderup, G. (2000). Corporate tax systems and cross country profit shifting. *Oxford Economic Papers*, *52*, 306–325.

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

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

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

Hindriks, J., & Nishimura, Y. (2018). *Taxing multinationals: The scope for enforcement cooperation* (CORE Discussion Paper, 2018/29).

Hindriks, J., Peralta, S., & Weber, S. (2008). Competing in taxes and investment under fiscal equalization. *Journal of Public Economics*, *92*, 2392–2402.

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

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

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, *13*, 324–340.

Hong, Q., & Smart, M. (2010). In praise of tax havens: International tax planning and foreign direct investment. *European Economic Review*, *54*, 82–95.

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

Johannesen, N. (2010). Imperfect tax competition for profits, asymmetric equilibrium and beneficial tax havens. *Journal of International Economics*, *81*, 253–264.

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

Keen, M., & Konrad, K. (2013). The theory of international tax competition and coordination. In A. Auerbach, R. Chetty, M. Feldstein, & E. Saez (Eds.), *Handbook of public economics* (Vol. 5, pp. 257–328). Amsterdam: North Holland.

Keen, M., & Slemrod, J. (2017). Optimal tax administration. *Journal of Public Economics*, *152*, 133–142.

Kempf, H., & Rota-Graziosi, G. (2010). Endogenizing leadership in tax competition. *Journal of Public Economics*, *94*, 768–776.

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

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

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

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.

Næss-Schmidt, P., Sunesen, E., Hansen, M., & Ali, D. (2012). *Tax elasticities of financial instruments, profits and remuneration* (Working Paper No. 30-2012). European Commission, Luxembourg.

Nielsen, S. B., Raimondos-Møller, P., & Schjelderup, G. (2008). Taxes and decision rights in multinationals. *Journal of Public Economic Theory*, *10*, 245–258.

Nielsen, S. B., Raimondos-Møller, P., & Schjelderup, G. (2010). Company taxation and tax spillovers: Separate accounting versus formula apportionment. *European Economic Review*, *54*, 121–132.

OECD. (2015). *BEPS 2015 Final Reports*. Organisation for Economic Co-operation and Development, Paris. https://www.oecd.org/ctp/beps-2015-final-reports.htm

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

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.

Swenson, D. (2001). Tax reforms and evidence of transfer pricing, *National Tax Journal*, *54*, 7–25.

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

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

How to cite this article: Hindriks J, Nishimura Y. Taxing multinationals: The scope for enforcement cooperation. *J Public Econ Theory*. 2021;1–23. https://doi.org/10.1111/jpet.12501

APPENDIX A

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

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

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

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

$$-t_1 + t_2 - 4\delta(e)(\tilde{\pi}_1^k - \pi_1^k) = 0. \tag{A1}$$

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

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

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

where we use (A1). These first-order conditions yield the reaction function $q_i^k = \left(\gamma_i - \beta q_i^{k'}\right)/(2\beta)$, 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 $\sum_{k=a,b}\left(\pi_1^k + \pi_2^k\right) = 1$, $\pi_1^k = \frac{1+\epsilon}{4}$ and $\pi_2^k = \frac{1-\epsilon}{4}(k = a, b)$. From (A1), we obtain: $\tilde{\pi}_1^k = \pi_1^k - \frac{t_1 - t_2}{4\delta(e)}$ and $\tilde{\pi}_2^k = \pi_2^k + \pi_1^k - \tilde{\pi}_1^k = \pi_2^k + \frac{t_1 - t_2}{4\delta(e)}(k = a, b)$. Therefore, the conclusion of the lemma holds. □

Proof of Lemma 2 and Lemma 3. Given e_j, country i 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 (9) and $\delta(e)$ takes the form of (2). The first-order conditions are given by:

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

Solving the system, we have $e_1^N(\rho,\epsilon)=\frac{1}{\eta}\left(\frac{3+\epsilon}{6}\right)^2\left(0.5\left(\frac{3-\epsilon}{3+\epsilon}\right)^{\frac{-2\rho}{2+\rho}}+0.5\right)^{\frac{1+\rho}{-\rho}}$, $e_2^N(\rho,\epsilon)=\frac{1}{\eta}\left(\frac{3-\epsilon}{6}\right)^2\left(0.5\left(\frac{3+\epsilon}{3-\epsilon}\right)^{\frac{-2\rho}{2+\rho}}+0.5\right)^{\frac{1+\rho}{-\rho}}$, and $\delta^N(\rho,\epsilon)=\frac{1}{36\eta}\left(0.5\,(3+\epsilon)^{\frac{-2\rho}{2+\rho}}+0.5\,(3-\epsilon)^{\frac{-2\rho}{2+\rho}}\right)^{\frac{2+\rho}{-\rho}}$. These expressions include the limit cases of $e_i^N(-1,\epsilon)=\frac{1}{\eta}\left(\frac{3+\epsilon_i}{6}\right)^2(i=1,2,\epsilon_1=\epsilon=-\epsilon_2)$, $\delta^N(-1,\epsilon)=\frac{9+\epsilon^2}{36\eta}$, $e_i^N(0,\epsilon)=\frac{(3+\epsilon_i)^{3/2}(3-\epsilon_i)^{1/2}}{36\eta}$, $\delta^N(0,\epsilon)=\frac{(3+\epsilon)(3-\epsilon)}{36\eta}$ and $e_1^N(\infty,\epsilon)=\frac{(3-\epsilon)^2}{36\eta}\frac{(3+\epsilon)^2}{\epsilon^2+9}=e_2^N(\infty,\epsilon)=\delta^N(\infty,\epsilon)$. 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 $x\equiv\left(\frac{3+\epsilon}{3-\epsilon}\right)^2\in[1,4)$ and $y\equiv\frac{\rho}{2+\rho}\in[-1,1)\backslash\{0\}$, we have:

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

Let $Z(x,y)\equiv((0.5x^{-y}+0.5)\ln(0.5x^{-y}+0.5)-(0.5+0.5y)x^{-y}\ln(x^{-y}))$. 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 $\ln x<\frac{(\ln z)^2}{\ln z-(z^{-1}+1)\ln(0.5z+0.5)}\equiv f(z)$. Indeed, $f(z)>\ln 4$ for all $z\in(0.25,4)\backslash\{1\}$. Therefore, in the range of $x\in(1,4)$ and $y\in[-1,1)\backslash\{0\}$, 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 $\rho=0$, $\frac{\partial e_2^N(\rho,\epsilon)}{\partial\rho}=\frac{1}{36\eta}(3-\epsilon)^2\frac{1}{8}x^{\frac{1}{4}}\ln x\left(1-\frac{1}{4}\ln x\right)>0$ for $x\in(1,4)$. Likewise, one obtains $\partial e_1^N(\rho,\epsilon)/\partial\rho<0$ for all $\rho\geq -1$ and $\epsilon>0$.

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

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

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

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 $\rho < -\frac{2}{3} \equiv \tilde{\rho}$ (i.e., when enforcement efforts are sufficiently substitutable) and negative if $\rho > -\frac{2}{3}$ (i.e., when enforcement efforts are sufficiently complementary). For $\rho = 0$, $\frac{\partial \delta^N(\rho, \epsilon)}{\partial \epsilon} = -\frac{\epsilon}{18\eta}$. For $\rho = -\frac{2}{3}$, $\delta^N\left(-\frac{2}{3}, \epsilon\right) = \frac{1}{4\eta}$ for all $\epsilon \geq 0$. □

Proof of Proposition 1. The first-order conditions of the joint welfare maximization (11) are $e_i^{-\rho-1}(0.5e_1^{-\rho} + 0.5e_2^{-\rho})^{\frac{1+\rho}{-\rho}}\left(\frac{9+\epsilon^2}{18}\right) - \eta e_i = 0 (i = 1, 2)$, from which we obtain $\hat{e}_1^N(\epsilon) = \hat{e}_2^N(\epsilon) = \frac{9+\epsilon^2}{18\eta}$. When $\rho = -1, (\hat{\delta}^N(\epsilon) - \delta^N(-1, \epsilon))\frac{(3-\epsilon)^2}{18} = \frac{(\epsilon^2+9)(3-\epsilon)^2}{36 \cdot 18\eta}$ is decreasing in ϵ, and $\frac{\eta}{2}\left(\left(\hat{e}_2^N(\epsilon)\right)^2 - \left(e_2^N(-1, \epsilon)\right)^2\right) = \frac{(3+\epsilon)^2(\epsilon^2-2\epsilon+9)}{24 \cdot 36\eta}$ is increasing in ϵ. Taking these together, $W_2^N(\hat{e}^N(\epsilon)) - W_2^N(e^N(\rho, \epsilon))$ in the text is 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 $\rho > -1$, $W_2^N(\hat{e}^N(\epsilon)) = \frac{\hat{\delta}(\epsilon)}{2}\left(\frac{3-\epsilon}{3}\right)^2 - \eta\frac{\left(\hat{e}_2^N(\epsilon)\right)^2}{2}$ is invariant with respect to ρ, whereas $W_2^N(e^N(\rho, \epsilon)) = \frac{\delta^N(\rho, \epsilon)}{2}\left(\frac{3-\epsilon}{3}\right)^2 - \eta\frac{\left(e_2^N(\rho, \epsilon)\right)^2}{2}$ decreases in ρ from Lemma 3. We numerically show that there exists an increasing function $\epsilon^N(\rho)$ such that $W_2^N(\hat{e}^N(\epsilon)) - W_2^N(e^N(\rho, \epsilon))$ is positive (negative) when $\epsilon < \epsilon^N(\rho) (\epsilon > \epsilon^N(\rho))$. $\epsilon^N(\rho)$ is shown in Figure 3: 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 3, we have $W_1^N(e^N(\rho, \epsilon)) < \frac{\delta^N(-1, \epsilon)}{2}\left(\frac{3+\epsilon}{3}\right)^2 - \eta\frac{\left(e_1^N(\infty, \epsilon)\right)^2}{2} \equiv w(\epsilon)$ for all ρ. $W_1^N(\hat{e}^N(\epsilon)) - w(\epsilon) = \frac{\epsilon^8 + 24\epsilon^7 + 36\epsilon^5(1-\epsilon) + 612\epsilon^5 + 486\epsilon^4 + 5832\epsilon^3 + 2916\epsilon(1-\epsilon) + 14580\epsilon + 6561}{2 \cdot 36^2\eta(\epsilon^2+9)^2} > 0$. Therefore, we have $W_1^N(e^N(\rho, \epsilon)) < w(\epsilon) < W_1^N(\hat{e}^N(\epsilon))$ for all ρ and ϵ. □

Proof of Lemma 4.

From (6), the tax reaction function of country 2 is $\arg\max_{t_2} R_2(t_2, t_1, e) \equiv t_2^r(t_1; e) = \delta(e)\left(\frac{1-\epsilon}{2}\right) + \frac{t_1}{2}$. We have $\frac{\partial t_2^r}{\partial t_1} \in (0, 1)$. The first-order condition of the tax leader is given by $\frac{\partial W_1}{\partial t_1} + \frac{\partial W_1}{\partial t_2}\frac{\partial t_2^r}{\partial t_1} = \frac{1+\epsilon}{2} - \frac{t_1 - t_2^r(t_1)}{2\delta(e)} - t_1\frac{1-(1/2)}{2\delta(e)} = 0$,[22] which yields $t_1^S(e) = \delta(e)\left(\frac{3+\epsilon}{2}\right)$ and $t_2^S(e) = \delta(e)\left(\frac{5-\epsilon}{4}\right)$. From (7), we have $t_1^S(e) - t_2^S(e) = \frac{\delta(e)}{4}(1+3\epsilon) > \frac{2\delta(e)\epsilon}{3} = t_1^N(e) - t_2^N(e)$. For $\tilde{\pi}_i^S \equiv \tilde{\pi}_i\left(t_1^S(e), t_2^S(e), e\right)$, from Lemma 1, we have $\tilde{\pi}_1^S = \frac{1+\epsilon}{2} - \frac{t_1^S(e) - t_2^S(e)}{2\delta(e)} = \frac{3+\epsilon}{8} < \frac{5-\epsilon}{8} = \tilde{\pi}_2^S$. From (9), $0 < R_1^S(e) - R_1^N(e) = \frac{\delta(e)(3+\epsilon)^2}{144} < R_2^S(e) - R_2^N(e) = \frac{\delta(e)(3+\epsilon)(27-7\epsilon)}{288}$, as in part (i) of the lemma.

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

Then, $R_1^S(e) - R_2^S(e) = \frac{\delta(e)}{32}(-7 + \epsilon^2 + 22\epsilon) \gtrless 0 \Leftrightarrow \epsilon \gtrless 8\sqrt{2} - 11 \equiv \hat{\epsilon} \approx 0.31370850$. For $\epsilon < \hat{\epsilon}$, $0 < R_2^S(e) - R_1^S(e) \lesseqgtr R_1^N(e) - R_2^N(e) \Leftrightarrow \epsilon \gtrless \tilde{\epsilon} \approx 0.16094072$. 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 (ii) of the lemma holds. □

Proof of Lemma 5.

From Lemma 4 and $\rho = -1$, $W_1^S(e_1, e_2) = \frac{0.5e_1 + 0.5e_2}{16}(3 + \epsilon)^2 - \eta\frac{(e_1)^2}{2}$. 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 2, $e_1^N(-1, \epsilon) = \frac{(3+\epsilon)^2}{36\eta} < e_1^S$ and $e_2^N(-1, \epsilon) = \frac{(3-\epsilon)^2}{36\eta} < e_2^S$. As in Lemma 4(ii), $| e_2^S - e_1^S | < e_1^N(-1, \epsilon) - e_2^N(-1, \epsilon)$ if and only if $\epsilon > \tilde{\epsilon}$. □

Proof of Proposition 2.

From Lemma 4, 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 $\hat{e}_1^S = \hat{e}_2^S = \frac{3\epsilon^2 + 2\epsilon + 43}{64\eta}$. For the tax leader, $R_1^S(\hat{e}^S) - R_1^S(e^S) = \frac{(3\epsilon^2 + 2\epsilon + 43)(3+\epsilon)^2}{64 \cdot 32\eta}$ is increasing in ϵ, and $\frac{\eta}{2}\left(\left(\hat{e}_1^S\right)^2 - \left(e_1^S\right)^2\right) = \frac{(5-\epsilon)^2(5\epsilon^2 + 14\epsilon + 61)}{2 \cdot 64^2\eta}$ is decreasing in ϵ. Also, at $\epsilon = 0$, $W_1^S(\hat{e}^S) - W_1^N(e^S) = \frac{23}{8192\eta} > 0$. Therefore, the tax leader unambiguously gains from cooperation.

For the tax follower, $R_2^S(\hat{e}^S) - R_2^S(e^S) = \frac{(3\epsilon^2 + 2\epsilon + 43)(5-\epsilon)^2}{64^2\eta}$ is decreasing in ϵ, and $\frac{\eta}{2}\left(\left(\hat{e}_2^S\right)^2 - \left(e_2^S\right)^2\right) = \frac{(3+\epsilon)^2(\epsilon^2 - 2\epsilon + 17)}{32^2\eta}$ is increasing in ϵ. Taking these together, $W_2^S(\hat{e}^S) - W_2^N(e^S)$ in the text 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 $\epsilon^S > \epsilon^N(-1)$ and $\epsilon^S < 1$. □