MULTINATIONAL TAXATION UNDER PRESSURE: THE ROLE OF TAX DEDUCTIBILITY

Xuyang Chen, Jean Hindriks

LIDAM Discussion Paper CORE
2023 / 13

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/core-discussion-papers.html

Multinational taxation under pressure:

the role of tax deductibility

Xuyang Chen[1] **and Jean Hindriks**[2]

May 10, 2023

Abstract

To address international profit shifting by multinational enterprises, a number of countries are broadening their tax bases and using turnover tax to secure corporate tax revenues. This paper investigates the impact of tax deductibility on multinationals' behavior and corporate taxes by considering a model of corporate tax competition with two countries and a tax haven. A profit tax (with either separate accounting or formula apportionment) giving deductions for all costs can preserve production efficiency at the expense of tax base erosion. In contrast, a turnover tax with no deductibility of costs can eliminate profit shifting incentives at the expense of production distortion. We show that profit tax with formula apportionment dominates the other two tax regimes when the output elasticity is high and tax capacity is intermediate. In other cases, turnover tax (profit tax with separate accounting) can dominate under low (high) tax capacity. We then analyze the general case to allow for partial tax deductibility and revisit the Diamond-Mirrlees production efficiency theorem in the profit shifting context. We show that a lower deductibility can yield larger tax revenue under sufficiently low tax capacity. Alternatively, tax revenue strictly increases with tax deductibility under high tax capacity, low marginal production cost or high output elasticity.

JEL classification: H25, H71, H73, F23, D24
Keywords: Corporate taxes, Profit shifting, Tax haven, Tax deductibility, Tax competition, Digital economy, Extractive sector

1.Introduction

[1] Center for Operations Research and Econometrics (CORE/LIDAM), UCLouvain, Belgium. xuyang.chen@uclouvain.be. Xuyang Chen gratefully acknowledges financial support from Fédération-Wallonie-Bruxelles through "Action de recherche concertée" (Grant No. 19/24-101).

[2] Center for Operations Research and Econometrics (CORE/LIDAM), UCLouvain, Belgium. jean.hindriks@uclouvain.be.

Multinational enterprises (MNEs) can avoid taxes by shifting profits from high-tax countries to low-tax countries. A number of studies document that international profit shifting causes considerable losses of tax revenue for both OECD and developing countries (e.g., Crivelli, de Mooij and Keen, 2016; Clausing, 2016; Davies et al., 2018; Bilicka, 2019; Tørsløv et al., 2022). Base erosion and profit shifting (BEPS) persists and raises problems for taxation in various sectors of the economy. For instance, the operations of digital companies rely heavily on the use of intangible assets, while the lack of market parallels for firm-specific intangibles makes it difficult for tax authorities to determine the arm's length prices. This undermines countries' tax capacity and provides opportunities for MNEs to shift profits from production affiliates in high-tax countries to intangibles-holding low-tax affiliates by overstating the true transfer price for royalties and licenses (see Dischinger and Riedel, 2011; the Gaspar Committee, 2014; Zucman, 2014; European Commission, 2017; OECD, 2018). In addition to the digital sector, MNEs in traditional "bricks and mortar" industries and extractive sectors also engage in aggressive tax planning including transfer pricing and internal debt (see ActionAid, 2012, 2015; Kleinbard, 2013). In particular, Albertin et al. (2021) argue that the corporate tax revenue losses in African countries are between $470 million and $730 million per year on average due to mining MNEs' profit shifting.

In order to protect tax bases and secure tax revenues, countries are limiting the tax deductibility and using sectoral turnover taxes targeted at gross revenues. Specifically, over the last few years, many countries around the world have proposed, and in some cases already implemented digital services taxes (DSTs) to tax large businesses based on their revenues from certain digital services provided to domestic consumers (see OECD, 2018; Bunn et al., 2020). In particular, the European Commission proposed an interim DST on gross revenues in the European Union (EU) from specific digital services characterized by user value creation. The proposed single rate across the EU is at 3% (see European Commission, 2018; Klein et al., 2022). In extractive sectors, most countries' fiscal regimes are a combination of corporate income taxes and royalties. Royalties are payment made by companies to governments for the extraction of finite resources. They are usually levied on the gross value of the product and so are akin to turnover taxes. Moreover, nine of the 15 resource intensive economies in Africa implement alternative minimum taxes (AMTs) to ensure some corporate taxes are paid each year (see Albertin et al., 2021). Under AMTs, companies are subject to the corporate income tax or a tax on turnover, whichever is higher. The rationale for using turnover taxes is that it becomes less beneficial for MNEs to inflate costs in high-tax affiliates for profit shifting purpose when countries reduce the deductibility of costs. Compared to a pure profit tax which falls on economic rents and whose features are identified in many countries' tax systems (see Ernst & Young, 2015), a broader tax base with limited tax deductibility is less vulnerable to profit shifting strategies and is more difficult to avoid. This is similar to Best et al.'s (2015) argument that broadening tax bases can lead to less tax evasion in development countries. On the other hand, turnover taxes introduce distortions to production and render real output lower than the efficient level. The magnitude of the distortions is related to the specific production technologies used by MNEs and thus is likely to vary across sectors. For instance, the emerging technological trends – including cloud computing, the Internet of Things, advanced robotics, and artificial intelligence – are affecting the digital economy, which reduce marginal production costs and boost productivity (see the Gaspar Committee, 2014; OECD, 2018). These technological improvements can exacerbate the production inefficiency incurred by turnover taxation and affect the international tax design in unexpected directions.

This paper takes into account both MNEs' production technology and countries' tax capacity to investigate the impact of tax deductibility on MNEs' behavior and countries' corporate taxes in international settings. In this regard, our work can be related to Agrawal and Wildasin (2020), who study the effect of transaction technologies for consumers – captured by the cost of online commerce relative to traditional commerce – on countries' commodity taxes. In contrast, we analyze the implications of MNEs' production technologies – reflected by the marginal production cost and output elasticity – for the choice of different tax regimes. Formally, we set up a tax competition model, in which two countries compete in corporate tax rates to maximize their tax revenues with a tax haven levying zero tax rate. In response to international tax-rate differentials, the MNE may distort the intra-firm transfer prices to lower its tax liability. This is different from Best et al. (2015), who focus on the optimal tax design in a single-country framework with tax evasion. They analyze a conventional Ramsey-type tax problem where a country chooses both tax rate and tax base parameter to maximize a firm's after-tax profit subject to an exogenous tax revenue constraint. Our model starts from two polar cases in which all costs in the affiliates of two countries are either fully tax deductible (a pure profit tax) or fully non-deductible (a turnover tax). According to two alternative principles in taxing MNEs, profit taxation can be further divided into two different tax regimes: a profit tax with separate accounting (SA) and with formula apportionment (FA). While the prevailing tax principle at the international level is SA, several countries – like the US, Canada and Germany – are applying the FA principle at the national level. The main advantage of FA over SA is that the profit shifting incentives between FA countries are abolished due to tax bases consolidation. FA has been proposed by the European Commission as a replacement for existing corporate tax systems in the EU to tackle BEPS (see European Commission, 2016). A substantial strand of literature on public finance has analyzed the implications of a switch between SA and FA (e.g., Kind et al., 2005; Nielsen, 2010; Eichner and Runkel, 2011; Gresik, 2016; Mardan and Stimmelmayr, 2018). In this regard, our work is closely related to Riedel and Runkel (2007), who investigate a tax system in which two countries form a FA union while a third country still sticks to SA (i.e., water's edge regulation). In our model, the tax haven also stays outside the FA union so as to attract mobile profits from FA countries. However, unlike Riedel and Runkel (2007), whose main focus is on the fiscal externalities under different tax systems, we explore the tax revenue consequences of a tax regime shift. Our paper also relates to Auerbach and Devereux (2018), who analyze different forms of cash-flow taxation, including a cash-flow tax apportioned by sales. In their work, a cash-flow tax – which taxes all revenues and gives deductions for all costs – is essentially a pure profit tax, and they focus on the location (i.e., a source or destination base) of the tax. In our model, the corporate income is taxed on an origin (i.e., source) basis, which is a standard principle of the existing international tax system, and our focus is on the deductibility of costs.

As we will show, a profit tax with either SA or FA can preserve efficient production at equilibrium, whereas countries' tax bases are eroded by the MNE's profit shifting. In particular, although a profit tax with FA eliminates the profit shifting between FA countries, the MNE still shifts profit from the FA union to the tax haven. In contrast, turnover tax can remove profit shifting incentives but causes distortion of the real production. Besides, we identify a formula manipulation effect under profit tax with FA. It results from the MNE's output adjustment in response to tax rate changes, and imposes a downward pressure on FA countries' taxes. A higher tax capacity increases the revenue efficiency under profit taxation, and also makes the formula manipulation effect stronger. For a smaller marginal production cost or a higher output elasticity, the turnover tax incurs a larger production efficiency loss, and the formula manipulation effect also increases. The tax revenue comparisons of the three tax regimes are affected by the production technology and tax capacity through these mechanisms. We derive conditions under which one tax regime dominates the other two in terms of the equilibrium tax revenue. Furthermore, we relate our results to the international tax design in practice, showing that the choice of the tax regimes for highly digitalized businesses and for the mining industry can be very different.

We then study the general case of tax deductibility where countries' tax bases can range from a pure profit tax base to a turnover base. We show that contingent on the tax capacity and production technology, the deductibility affects equilibrium tax revenues differently. Specifically, a lower deductibility can lead to larger tax revenue when countries' tax capacity is very low. The insight is that profit shifting is severe and revenue efficiency is the first-order concern under low tax capacity. Then a smaller deductibility is beneficial, since it makes profit shifting less tax sensitive. Alternatively, under high tax capacity, low marginal production cost or large output elasticity, the benefit from increasing production efficiency is first-order, and equilibrium revenue is strictly increasing in tax deductibility. In these cases, a pure profit tax preserving full production efficiency is optimal, even though some profits go untaxed due to profit shifting. Therefore, we extend the Diamond-Mirrlees (DM) theorem (Diamond and

Mirrlees, 1971) – which states that tax systems should maintain full production efficiency – to the context of international tax competition and profit shifting. This can relate our paper to Keen and Wildasin (2004). Using a standard competitive trade model, they reveal that in many-country settings Pareto-efficient international taxation may require production inefficiency and establish conditions under which the DM theorem still holds. In their model, production inefficiency may be desirable due to the distinctness of national budget constraints. In our paper, countries may benefit from a production-inefficient tax policy due to the aggressive profit shifting. Besides, numerical simulations further illustrate that the optimal deductibility rate will increase for higher tax capacity and output elasticity, and lower marginal production cost. The insight is that those parameter changes tilts the trade-off between production and revenue efficiency towards the former, making the optimal tax base closer to a pure profit tax base.

The paper is organized as follows. Section 2 sets up a model of international tax competition with two countries and a tax haven. Section 3 characterizes equilibrium outcomes in different tax regimes – a pure profit tax with either SA or FA and a turnover tax. Section 4 compares the three tax regimes in terms of equilibrium tax revenues. In Section 5, we go beyond the polar cases – profit and turnover taxation – to investigate the general case of tax deductibility. Section 6 concludes with a few policy remarks.

2. The model

Consider an economy with two identical countries labelled by 1, 2 and a tax haven indexed by 3. Each country hosts an affiliate of a representative multinational enterprise (MNE). The affiliates in two countries produces a homogeneous good, which is sold in the world market at a price normalized to unity. The output of the affiliate in country i ($i=1,2$) is denoted by y_i, and the production cost function is assumed to take the convex-linear form: $c(y_i) := \frac{\alpha}{k} y_i^k + \beta y_i$ with $\alpha > 0$, $0 \le \beta < 1$ and $k > 1$. Both α and β measure the magnitude of the marginal production cost. According to Palmer (2003), k captures the relative convexity of the marginal production cost function. Moreover, as we will show, k is inversely related to the elasticity of real output. For simplicity, in the paper we mainly focus on technological parameters α and k in the convex component of $c(y_i)$. In Appendix C, we prove that our result for the comparison of profit and turnover taxation also holds with the general form of production cost function. The resource endowment and market demand in the tax haven is negligible such that the MNE has no real production there.

The MNE can shift profits between affiliates in order to minimize its tax liability. In practice, the MNE can locate intangible assets in low-tax affiliates, which then charge transfer prices to affiliates in high-tax countries for the use of intangibles. The arm's length prices for the intangible assets are normalized to zero. Denote by g_{ij} the intra-firm transfer price between affiliates i and j ($i, j \in \{1,2,3\}$; $i \neq j$). Specifically, $g_{ij} > 0$ ($g_{ij} < 0$) represents that the transfer price is paid by (to) affiliate i, and hence profit is shifted to (from) country j. Notably, profit shifting does not change the MNE's overall profits, as reflected by:

$$g_{ij} + g_{ji} = 0. \tag{1}$$

The profit shifting between affiliates i and j is costly for the MNE and incurs a convex concealment cost with the form $h(g_{ij}) := \frac{\delta}{2} g_{ij}^2$ ($1 \leq i < j \leq 3$), where parameter $\delta > 0$ captures countries' tax capacity (see, e.g., Kind et al., 2005; Devereux et al., 2008; Mardan and Stimmelmayr, 2018). The statutory tax rate in country i ($i = 1,2$) is denoted by t_i. For simplicity, we assume that the tax haven imposes no taxes at all on profits shifted there, which is a usual assumption in literature on international profit shifting (e.g., Slemrod and Wilson, 2009; Johannesen, 2010; Gumpert et al., 2016). An alternative to the exogenous zero rate is to assume that the tax haven maximizes its tax revenue with shifted profits being its only tax base (e.g., Krautheim and Schmidt-Eisenlohr, 2011). Our findings remain qualitatively unchanged if we endogenize the tax rate of the tax haven.

Each affiliate's net-of-tax profit is given by:

$$\begin{aligned} \pi_i &= y_i - c(y_i) - \sum_{1 \leq j \leq 3, j \neq i} g_{ij} - t_i \pi_i^t, \quad i = 1,2, \\ \pi_3 &= \sum_{i=1}^{2} g_{i3}, \end{aligned} \tag{2}$$

where π_i^t is the MNE's taxable income in country i which depends on the tax regime applied by two countries and will be presented in the next sections.

We consider a two-stage tax competition game specified as follows. In the first stage, two countries choose tax rates simultaneously and non-cooperatively to maximize their own tax revenues. In the second stage, the MNE chooses outputs and intra-firm transfer prices to maximize total after-tax profits.

3. Corporate tax regimes

In this section, we derive equilibrium outcomes under profit and turnover taxation. Profit taxation can be further divided into two different tax regimes, according to two alternative principles in taxing MNEs (i.e., separate accounting and formula apportionment). We analyze each corporate tax regime separately.

3.1. Profit tax with separate accounting

Suppose that two countries use SA. Under SA, profit is taxed in the country where the MNE declares it. The taxable profit of the MNE in each country reads:

$$\pi_i^t = y_i - c(y_i) - \sum_{1\le j\le 3, j\ne i} g_{ij}, \quad i = 1,2. \tag{3}$$

Given (2) and (3), summing up the three affiliates' net-of-tax profits minus the concealment costs lead to the total after-tax profit of the MNE:

$$\Pi^{SA} = \sum_{i=1}^{2}\left[(1-t_i)\left(y_i - c(y_i) - \sum_{1\le j\le 3, j\ne i} g_{ij}\right)\right] + \sum_{i=1}^{2} g_{i3} - \sum_{1\le i<j\le 3}\frac{\delta}{2}g_{ij}^2. \tag{4}$$

The MNE chooses output y_i and the amount of profit shifting g_{ij} ($1 \le i < j \le 3$) to maximize (4) subject to (1). Solving the first-order conditions of the MNE's profit maximization gives:

$$c'(y_i) = 1, \quad i = 1,2, \tag{5}$$

$$g_{ij} = \frac{t_i - t_j}{\delta}, \quad i, j = 1,2; \quad i \ne j, \qquad g_{i3} = \frac{t_i}{\delta}, \quad i = 1,2. \tag{6}$$

From (5), the optimal output chosen by two affiliates is independent of the tax system under SA. Two countries achieve *full* production efficiency with the efficient output choice $\bar{y} = \left(\frac{1-\beta}{\alpha}\right)^{\frac{1}{k-1}}$. We denote by $\bar{\pi} := \bar{y} - c(\bar{y}) = \frac{(1-\beta)(k-1)}{k}\bar{y}$ the actual profit of each affiliate under profit tax. It is readily verified that both $\bar{y}$ and $\bar{\pi}$ are strictly decreasing in α for all $\alpha > 0$ and in k when $\alpha \le \frac{1-\beta}{e}$. In this paper, we keep the assumption $0 < \alpha \le \frac{1-\beta}{e}$, while it is only used for Proposition 1(ii) and Proposition 2(iii) with respect to parameter k.

Given (3), (5) and (6), the tax base loss from a marginal increase in tax rate is $\frac{\partial}{\partial t_i}\sum_{1\le j\le 3, j\ne i} g_{ij} = \frac{2}{\delta}$. This tax sensitivity of profit shifting captures the revenue efficiency under SA. A higher tax capacity makes profit shifting less sensitive to tax rate changes and indicates a gain in revenue efficiency.

In the first stage, country i takes as given the tax rate of country j and sets tax rate t_i to maximize its tax revenue $R_i(t_i, t_j) = t_i \pi_i^t$, where π_i^t is given by (3). In doing so, it takes into account the MNE's profit maximizing solutions which depend on t_i and t_j according to (5) and (6). Solving the first-order conditions of tax revenue maximization, we can get the Nash equilibrium tax rate and revenue of each country:

$$\tilde{t}(\delta) = \min\left\{\frac{\bar{\pi}\delta}{3}, 1\right\}, \tag{7}$$

$$\tilde{R}(\delta) = \begin{cases} \dfrac{2\bar{\pi}^2\delta}{9} & \text{if } 0 < \delta < \dfrac{3}{\bar{\pi}} \\ \bar{\pi} - \dfrac{1}{\delta} & \text{if } \delta \geq \dfrac{3}{\bar{\pi}} \end{cases}. \tag{8}$$

Given (7) and (8), it is straightforward to establish the following:

Lemma 1. *Under SA, (i) there exists a unique symmetric Nash equilibrium; (ii) each country's equilibrium tax rate* $\tilde{t}$ *is strictly increasing in tax capacity when* $\delta < \frac{3}{\bar{\pi}}$ *, and attains 100% when* $\delta \geq \frac{3}{\bar{\pi}}$*; (iii) the equilibrium tax revenue* $\tilde{R}$ *is strictly increasing in the tax capacity.*

Intuitively, a higher tax capacity increases MNE's concealment costs of profit shifting. This reduces the tax sensitivity of profit shifting and increases the revenue efficiency under SA, which enables each country to tax more and gain higher revenues.

3.2. Profit tax with formula apportionment

Consider that country 1 and 2 form a FA union. In the FA union, two countries' tax bases are consolidated first and then assigned to each country based on the MNE's sales in each country relative to its global sales revenue. The tax haven will stay outside the union and sticks to SA, in order to attract the MNE's profits from the FA countries (i.e., a water's edge regulation). In practice, the tax haven attracts extensive foreign capital to its financial sector, which is then invested for profits. Hines (2005) documents that tax havens have enjoyed very rapid economic growth due to large volumes of foreign investment. The MNE's sales share in country i is $\lambda_i := \frac{y_i}{\sum_{j=1}^{2} y_j}$. Given (1) and (3), the consolidated tax base in the FA union is:

$$\pi^c := \sum_{i=1}^{2}[y_i - c(y_i) - \sum_{1\le j\le 3, j\ne i} g_{ij}] = \sum_{i=1}^{2}[y_i - c(y_i) - g_{i3}]. \tag{9}$$

The MNE's taxable income in country i reads:

$$\pi_i^t = \lambda_i \pi^c, \quad i = 1,2. \tag{10}$$

Given (1), (2), (9) and (10), the MNE's total after-tax profit under FA reads:

$$\Pi^{FA} = (1-\tau)\pi^c + \sum_{i=1}^{2} g_{i3} - \sum_{1\le i<j\le 3} \frac{\delta}{2} g_{ij}^2,$$

where $\tau := \sum_{i=1}^{2} \lambda_i t_i$ is the effective tax rate within the FA union.

The first-order conditions of the MNE's profit maximization are:

$$\frac{\partial \Pi^{FA}}{\partial g_{12}} = -\delta g_{12} = 0, \tag{11}$$

$$\frac{\partial \Pi^{FA}}{\partial g_{i3}} = \tau - \delta g_{i3} = 0, \quad i = 1,2, \tag{12}$$

$$\frac{\partial \Pi^{FA}}{\partial y_i} = -\frac{\partial \tau}{\partial y_i}\pi^c + (1-\tau)\frac{\partial \pi^c}{\partial y_i} = \frac{(\tau - t_i)\pi^c}{\sum_{i=1}^{2} y_i} + (1-\tau)[1 - c'(y_i)] = 0, \quad i = 1,2. \tag{13}$$

Because of the tax base consolidation within the FA union, (11) indicates that the profit shifting between FA countries is eliminated. Nevertheless, the MNE's incentive of profit shifting still exists, since the tax haven always undercuts the FA countries. As shown by (12), the MNE can transfer profits from the FA union to the tax haven to minimize the total tax liability. In the symmetric situation with identical tax rates $t_i = t$ ($i = 1,2$), we have $\tau = t$. Then it follows from (13) that $y_i = \bar{y}$ with $c'(\bar{y}) = 1$. This implies that the two FA countries can preserve *full* production efficiency at the symmetric equilibrium.

Totally differentiating (12) and (13), together with the symmetry property, yields the comparative statics under identical tax rate $t_i = t$ ($i = 1,2$):

$$\frac{\partial g_{i3}}{\partial t_i} = \frac{\partial g_{j3}}{\partial t_i} = \frac{1}{2\delta} > 0, \quad i, j = 1,2, \tag{14}$$

$$\frac{\partial y_i}{\partial t_i} = -\frac{\bar{\pi} - t/\delta}{2(1-t)\bar{y}c''(\bar{y})} < 0, \quad i = 1,2, \tag{15}$$

$$\frac{\partial y_j}{\partial t_i} = \frac{\bar{\pi} - t/\delta}{2(1-t)\bar{y}c''(\bar{y})} > 0\,, \quad i,j = 1,2; \quad i \neq j\,. \tag{16}$$

Under FA, the MNE's response to one country's tax increase is twofold. Firstly, as shown by (14), the MNE shifts more profits from each FA country to the tax haven, since the unilateral tax increase raises the effective tax rate in the FA union. Secondly, as shown in (15) and (16), the MNE manipulates the sales formula by favorably adjusting the output in two FA countries. It increases (decreases) output in the low-tax (high-tax) country so as to increase the share of the consolidated tax base assigned to the low-tax country and reduce the tax burden. Furthermore, a higher tax capacity reduces the profit shifted to the tax haven and yields a larger consolidated tax base subject to taxation within the FA union. This reinforces the MNE's output adjustment in response to the tax rate changes. Similarly, a higher productivity of the MNE (i.e., a smaller α or k) also leads to a larger consolidated tax base, which makes each affiliate's output more tax sensitive.

In the first stage, country i takes as given the tax rate of country j and sets tax rate t_i to maximize its tax revenue $R_i = t_i \lambda_i \pi^c$, where π^c is given by (9). In doing so, it takes into account the MNE's production and profit shifting choices which rely on countries' tax rates according to (11) – (13). Following the previous literature, we focus on the symmetric Nash equilibrium in which two FA countries set equal taxes $\hat{t}$. Solving the first-order conditions of tax revenue maximization and then using the symmetry property, we obtain the following:

$$\frac{\partial R_i}{\partial t_i} = \bar{\pi} - \frac{3\hat{t}}{2\delta} - \frac{\hat{t}(\bar{\pi} - \frac{\hat{t}}{\delta})^2}{2(1-\hat{t})\bar{y}^2 c''(\bar{y})} = 0\,. \tag{17}$$

(17) determines two countries' equilibrium tax rates. Due to the asymmetry between FA countries and the tax haven, the solution to (17) has no simple expressions. Despite this complexity, we can establish the following:

Lemma 2. *Under FA, (i) there exists a unique symmetric Nash equilibrium with $\hat{t} \in (0,1)$; (ii) each country's equilibrium tax $\hat{t}$ is strictly increasing (decreasing) in tax capacity when $0 < \delta < \delta^*$ ($\delta > \delta^*$), where $\delta^* := \frac{27k}{(4+9k)\bar{\pi}}$ is strictly increasing in α and k; (iii) the equilibrium tax revenue $\hat{R}$ is strictly increasing in the tax capacity.*

Proof. See Appendix A1.

Different from SA where equilibrium taxes are increasing in the tax capacity, under FA there is a bell-shaped relationship between equilibrium taxes and tax capacity. To understand the rationale of this result, using (14) – (16) and the symmetry condition, we compute the loss of one country's tax base from increasing its own tax rate: $-\left.\frac{\partial \pi_i^t}{\partial t_i}\right|_{t_i=t_j=t} = \lambda_i \cdot \frac{\partial}{\partial t_i}\sum_{j=1}^{2} g_{j3} - \frac{\partial \lambda_i}{\partial t_i}\cdot \pi^c$. The first term on the right-hand side is the tax sensitivity of profit shifting ($\lambda_i \cdot \frac{\partial}{\partial t_i}\sum_{j=1}^{2} g_{j3} = \frac{1}{2\delta}$), which reflects the taxable income loss due to profit shifting and captures the revenue efficiency under FA. The second term captures the "formula manipulation effect" ($-\frac{\partial \lambda_i}{\partial t_i}\cdot \pi^c = \frac{(\bar{\pi} - t/\delta)^2}{2(1-t)k\bar{\pi}}$), which reflects the taxable income loss due to the MNE's manipulation of sales share. Increasing tax capacity has two opposite effects on the tax rate. On one hand, it reduces the tax sensitivity of profit shifting and induces each FA country to increase tax. On the other hand, recall that with higher tax capacity affiliates' output reacts more sensitive to the tax changes. This increases the formula manipulation effect and induces each FA country to tax less. Under a low (high) tax capacity, profit shifting (formula manipulation) is the main channel through which the MNE minimizes its tax liability, implying that the positive (negative) effect of tax capacity dominates. Moreover, for a smaller α or k, each affiliate's output is more tax sensitive due to a larger consolidated tax base at stake, which makes the formula manipulation effect stronger. Consequently, the threshold δ^* declines such that the range within which equilibrium tax rate decreases with tax capacity is wider. For the tax revenue, a higher tax capacity reduces the profit shifted to the tax haven. But it also involves a stronger formula manipulation effect, exerting a negative effect on FA countries' taxes. Lemma 2(iii) reveals that the increase in the equilibrium tax base outweighs the possible tax rate reduction due to formula manipulation.

3.3. Turnover taxation

Consider now two countries apply a turnover tax regime. Then the MNE's taxable income in country i is the gross revenue (i.e., turnover) of affiliate i:

$$\pi_i^t = y_i, \quad i = 1,2. \tag{18}$$

Given (1), (2) and (18), the MNE's total after-tax profit under turnover tax amounts to:

$$\Pi^T = \sum_{i=1}^{2}[(1-t_i)y_i - c(y_i)] - \sum_{1\le i<j\le 3}\frac{\delta}{2}g_{ij}^2.$$

The MNE chooses output y_i and profit shifting g_{ij} ($1 \le i < j \le 3$) to maximize Π^T. The tax base in each country is the affiliate's real output, implying that the MNE's tax burden is irrelevant to profit shifting activities. This leads to:

$$g_{ij} = 0, \quad 1 \le i < j \le 3. \tag{19}$$

As shown by (19), the MNE's profit shifting is completely removed, and two countries can achieve *full* revenue efficiency under turnover tax, irrespectively of the international tax-rate differentials.

The output in each country is given by the first-order condition:

$$\frac{\partial \Pi^T}{\partial y_i} = 1 - t_i - c'(y_i) = 0, \quad i = 1, 2. \tag{20}$$

In contrast to profit tax that yields efficient output in each country, the turnover tax undermines the production efficiency, rendering output level smaller than $\overline{y}$.

Under turnover taxation, country i's tax revenue is simply given by: $R_i = t_i y_i$. In the first stage, country i chooses tax rate t_i to maximize its revenue subject to (20). Solving the first-order condition of the maximization problem yields the equilibrium turnover tax rate:

$$t^T = \frac{(1-\beta)(k-1)}{k}. \tag{21}$$

Given (20) and the specification of $c(y_i)$, following Best et al. (2015), the output elasticity is: $\varepsilon_i := \dfrac{dy_i / y_i}{d(1-t_i)/(1-t_i)} = \dfrac{1-t_i}{(k-1)(1-t_i-\beta)}$, which varies inversely with parameter k. So, a smaller k means a more elastic tax base in two countries, which induces them to tax less. The equilibrium tax rate varies from zero as $k \to 1$ to $1-\beta$ as $k \to +\infty$.

Using (20) and (21) yields the equilibrium output and revenue of each country:

$$\underline{y} = k^{\frac{1}{1-k}} \overline{y}, \tag{22}$$

$$R^T = k^{\frac{1}{1-k}} \overline{\pi}. \tag{23}$$

In (22) and (23), the deviation of output and tax revenue from the efficient levels is reflected by a discount factor $k^{\frac{1}{1-k}} < 1$. It is readily verified that $k^{\frac{1}{1-k}}$ is strictly increasing with k, implying that the tax distortion deteriorates for a smaller k. From

(22), the output gap between profit and turnover tax is $\overline{y}-\underline{y}=\left(1-k^{\frac{1}{1-k}}\right)\overline{y}$, which captures production efficiency loss incurred by turnover tax. Either a smaller marginal production cost (smaller α) or a larger output elasticity (smaller k) increases the loss of production efficiency under turnover tax.

We summarize the main result in the lemma below.

Lemma 3. *Under turnover tax, (i) profit shifting is completely eliminated and two countries can achieve full revenue efficiency; (ii) each country's equilibrium tax rate* t^T *and revenue* R^T *are independent of the tax capacity.*

4. Comparison of the tax regimes

4.1. Profit versus turnover taxation

As we have shown, a tax regime shift between profit and turnover taxation will change countries' tax rates and the MNE's decision on real production and profit shifting. In the following, we compare profit and turnover taxation from the perspective of countries' equilibrium revenues.

Proposition 1. *(i) There exist tax capacity thresholds* δ^{SA} *and* δ^{FA} *such that turnover tax is superior (inferior) to SA if* $\delta<\delta^{SA}$ $(\delta>\delta^{SA})$ *and that turnover tax is superior (inferior) to FA if* $\delta<\delta^{FA}$ $(\delta>\delta^{FA})$*. In particular, FA is always superior to turnover taxation in the absence of the tax haven. (ii) Both* δ^{SA} *and* δ^{FA} *strictly increase with technological parameters* α *and* k.

Proof. See Appendix A2.

Profit tax preserves full production efficiency at equilibrium, while the revenue efficiency is strictly increasing in the tax capacity. In contrast, with turnover tax countries achieve full revenue efficiency at the expense of inefficient production. The MNE's profit shifting incentive is completely eliminated, and tax capacity is irrelevant under turnover tax. It is then intuitive that a high (low) tax capacity is favorable to SA (turnover taxation). Under FA, there is an additional formula manipulation effect that negatively affects countries' taxes. Proposition 1(i) reveals that with a high tax capacity, the gain in production efficiency can overcompensate both formula manipulation effect and the revenue efficiency loss when countries switch from turnover tax to FA. Moreover, in an economy without tax haven, no profit flows out from the FA union, and two countries can preserve full revenue efficiency as under turnover tax. This is equivalent to the case of an infinitely high tax capacity, in which FA always dominates turnover tax. For part (ii), either a smaller α or k means a larger loss of production efficiency under turnover tax and hence favors profit tax. This implies a smaller threshold δ^{SA} so that the scope within which SA dominates turnover tax increases. Moreover, for a smaller α or k, our finding implies that the larger gain in production efficiency outweighs the stronger formula manipulation effect when switching from turnover tax to FA, so that threshold δ^{FA} decreases.

It is worth noting that when $k > 4.939$ and $\frac{3}{\pi} \leq \delta < \delta^{SA}$, a distortionary turnover tax still dominates a production-efficient profit tax, even though countries can set a tax rate of 100 per cent under profit tax. This result stands in sharp contrast to the classic optimal tax theory in the single-country framework which claims that full production efficiency is desirable if a 100 per cent profit tax can be levied (e.g., Stiglitz and Dasgupta, 1971; Munk, 1980). The reason is that in the context of profit shifting, the MNE always shifts profits from each country to the tax haven so that profit tax does not fall on real profits.

4.2. Comparing the three tax regimes

In this subsection, we provide the tax ranking of the three tax regimes (i.e., SA, FA and turnover tax). First, we compare SA and FA:

Lemma 4. *Each country's equilibrium tax and revenue are higher (lower) under SA than under FA if* $\delta > \delta^*$ ($0 < \delta < \delta^*$).

Proof. See Appendix A3.

Because the profit shifting incentive within FA union is eliminated, profit shifting is less tax sensitive and the revenue efficiency is higher under FA than under SA. On the other hand, FA introduces a formula manipulation effect which imposes a downward pressure on countries' tax rates. With a low tax capacity, the profit shifting is severe and the revenue efficiency concern is strong so that FA is superior to SA. For a high tax capacity, profit shifting is minor and the revenue efficiency in both regimes is large. However, the formula manipulation effect under FA is strong, so that SA can lead to higher equilibrium taxes than FA.

We are now ready to identify the optimal choice among three tax regimes in terms of countries' equilibrium tax revenues.

Proposition 2. *The optimal choice among the three tax regimes depends on the tax capacity and the MNE's production technology in the following way:*

(i) *for* $1 < k < 3.089$ *, the optimal choice is turnover taxation when* $0 < \delta < \delta^{FA}$*; FA when* $\delta^{FA} < \delta < \delta^{*}$*; and SA when* $\delta > \delta^{*}$.

(ii) *for* $k \geq 3.089$*, the optimal choice is turnover taxation when* $0 < \delta < \delta^{SA}$ *and is SA when* $\delta > \delta^{SA}$.

(iii) *the scope within which SA (turnover taxation) is optimal strictly decreases (increases) with* α *and* k.

Proof. See Appendix A4.

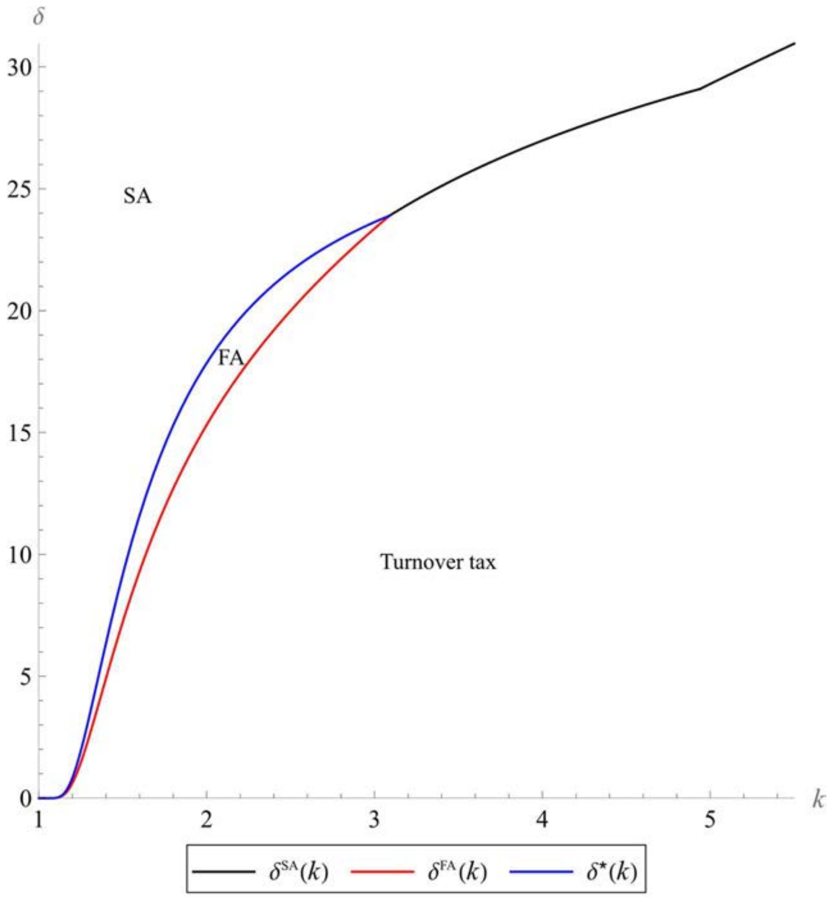


Fig. 1. The optimal choice among three tax regimes. Parameter values are specified as: $\alpha = \dfrac{2}{55}$ and $\beta = \dfrac{9}{10}$.

Fig. 1 illustrates how the optimal choice among the three tax systems depend on the configuration of k and δ. In the figure, $\delta^{SA}(k)$, $\delta^{FA}(k)$ and $\delta^{*}(k)$ all converge to zero as $k \to 1$. The three curves are upward-sloping and intersect at

$k = 3.089$ such that the (k, δ) space can be divided into three areas in terms of tax revenue rankings. FA is superior to the other two regimes only for high output elasticity (i.e., small k) and intermediate tax capacity. In other cases, SA (turnover tax) can dominate under high (low) tax capacity. The intuition underlying Proposition 2(i) and (ii) is as follows. When k is small, turnover tax incurs large loss of production efficiency due to the large output elasticity. Consequently, turnover tax can dominate only if the tax capacity is so low that revenue efficiency under profit tax is severely undermined by the MNE's profit shifting. When the tax capacity is intermediate, profit tax is superior to turnover tax. Moreover, FA is preferable to SA, since it can better cope with profit shifting and leads to higher revenue efficiency. For a sufficiently high tax capacity, the adverse effect of the formula manipulation under FA is strong, which makes SA preferable. When k is large, the output elasticity is small and the loss of production efficiency caused by turnover tax is limited. Therefore, profit tax is desirable only under a sufficiently high tax capacity. However, in such case SA is superior to FA due to the strong formula manipulation effect under FA. Hence, when k is large, countries only choose between turnover tax and SA. For part (iii), a smaller α or k favors SA, since it implies a larger production efficiency loss under turnover tax and a stronger formula manipulation effect under FA. When α or k rises, turnover tax causes a smaller production efficiency loss, which increases the scope within which it dominates profit taxation.

Our results have policy implications for the choice of tax regimes in practice. Notice that the pre-tax profit margin of the affiliate in each country under SA is: $\frac{\bar{y} - c(\bar{y})}{\bar{y}} = (1-\beta)(1-\frac{1}{k})$, which means that the pre-tax profit margin strictly increases with k. In the mining sector, both marginal production costs and profit margins are relatively high[3], which implies large values of α, β and k. Together with the low local capacity (i.e., small δ) of African tax authorities (see Albertin et al., 2021), the tax capacity and production technology in Africa's mining industry are likely to satisfy $k \geq 3.089$ and $0 < \delta < \delta^{SA}$. Then we can infer that it will be beneficial for African countries to switch from profit-based tax to turnover tax in the mining sector. This is in keeping with Keen and Mullins' (2017, p.34) argument that mining fiscal regimes should "tilt(ing) the balance between profit-related taxes and royalties further towards the latter than might otherwise be the case". By contrast, highly digitalized businesses display high volumes, low margins and very low marginal production cost (see the Gaspar Committee, 2014; OECD, 2018). We may therefore expect small values of α, β and k. Then Proposition 2(iii) suggests that for highly digitalized businesses a profit tax is likely to dominate a turnover tax and that the revenue-based DSTs should

[3] See https://pages.stern.nyu.edu/~adamodar/New_Home_Page/data.html.

be removed. A rigorous empirical analysis of the tax capacity and technological parameters is necessary to decide whether it is appropriate to ring-fence some sectors with industry-specific tax regime. We leave this for future research.

5. The general case of tax deductibility

We have so far analyzed the pure profit tax (under which costs are fully tax deductible) and turnover tax (under which costs are nondeductible). Now we depart from the two polar cases and analyze the general case of the deductibility. Following Best et al. (2015), denote by $\mu \in [0,1]$ the share of costs deductible from taxes. Then a smaller (larger) value of μ indicates a broader (narrower) tax base. Here we restrict attention to SA, which is the prevailing tax principle at the international level. The analysis of tax deductibility under FA is relegated to Appendix B.

When two countries and the tax haven apply SA, the MNE's taxable income in country i ($i=1,2$) reads:

$$\pi_i^t = y_i - \mu\left(c(y_i) + \sum_{1\le j\le 3, j\ne i} g_{ij}\right). \tag{24}$$

Given (1), (2) and (24), the MNE's total net-of-tax profit amounts to:

$$\Pi^{SA} = \sum_{i=1}^{2}(1-t_i)\left[y_i - \mu\left(c(y_i) + \sum_{1\le j\le 3, j\ne i} g_{ij}\right)\right] - (1-\mu)\sum_{i=1}^{2} c(y_i) + \mu\sum_{i=1}^{2} g_{i3} - \sum_{1\le i<j\le 3}\frac{\delta}{2}g_{ij}^2 . \tag{25}$$

In the second stage, the MNE chooses outputs and intra-firm transfer prices to maximize (25), subject to (1). Solving the first-order conditions of the profit maximization yields:

$$c'(y_i) = 1 - m_i, \quad i = 1,2, \tag{26}$$

$$g_{ij} = \frac{\mu(t_i - t_j)}{\delta}, \quad i,j = 1,2;\ i \ne j, \qquad g_{i3} = \frac{\mu t_i}{\delta}, \quad i = 1,2, \tag{27}$$

where $m_i := \dfrac{(1-\mu)t_i}{1-\mu t_i}$ is the effective marginal tax rate (EMTR) in country i.

From (27), the tax sensitivity of profit shifting is $\dfrac{\partial}{\partial t_i}\sum_{1\le j\le 3, j\ne i} g_{ij} = \dfrac{2\mu}{\delta}$. A higher tax deductibility makes profit shifting more sensitive to tax changes and indicates a loss of revenue efficiency.

Using (26), we can derive the following comparative statics:

$$\frac{\partial y_i}{\partial t_i}=-\frac{\partial m_i/\partial t_i}{c''(y_i)}=-\frac{1-\mu}{(1-t_i\mu)^2c''(y_i)}<0,\quad \forall\mu\in[0,1), \tag{28}$$

$$\frac{\partial y_i}{\partial \mu}=-\frac{\partial m_i/\partial \mu}{c''(y_i)}=\frac{(1-t_i)t_i}{(1-t_i\mu)^2c''(y_i)}>0,\ \forall\mu\in[0,1). \tag{29}$$

As in Best et al. (2015), (28) shows that increasing one country's tax rate raises the EMTR and reduces real output in this country. From (29), a higher tax deductibility reduces the EMTR and increases real output, which indicates a gain in production efficiency.

In the first stage, country i chooses tax rate t_i to maximize its revenue $R_i(t_i,t_j)=t_i\pi_i^t$, where π_i^t is given by (24). In doing so, it takes into account the MNE's response to countries' tax rates as in (26) and (27). As previously, we restrict attention to the symmetric Nash equilibrium. Solving the first-order conditions of tax revenue maximization and then using the symmetry condition yields the following:

$$\frac{\partial R_i(t_i,t_j)}{\partial t_i}=y-\mu c(y)-\frac{t(1-\mu)^2}{(1-\mu t)^3c''(y)}-\frac{3\mu^2 t}{\delta}=0, \tag{30}$$

where y is determined by (26) (with country subscripts dropped due to symmetry).

Given the specification of production cost function, solving (30) yields the Nash equilibrium tax rate $\tilde{t}(\delta,\mu)$ under SA. Due to the inherent complexity under partial deductibility, this equation has no closed-form solution for $\mu\in(0,1)$. However, we can establish the existence and uniqueness of the symmetric equilibrium (see Appendix A5). Besides, (26) implies that each country's equilibrium output $\tilde{y}$ directly depends on the equilibrium tax and deductibility parameter, i.e., $\tilde{y}=\tilde{y}(\tilde{t},\mu)$. In Appendix A6, we show that $\forall\mu\in(0,1)$, $\tilde{y}$ lies on the interval $(\underline{y},\overline{y})$. Given (24) and (27), each country's equilibrium revenue $\tilde{R}(\delta,\mu)$ is expressed as:

$$\tilde{R}(\delta,\mu)=\tilde{t}\,\tilde{\pi}^t(\tilde{t},\delta,\mu), \tag{31}$$

where $\tilde{\pi}^t(\tilde{t},\delta,\mu):=\tilde{y}(\tilde{t},\mu)-\mu\left[c(\tilde{y}(\tilde{t},\mu))+\tilde{g}(\tilde{t},\delta,\mu)\right]$ is the equilibrium tax base, with $\tilde{g}(\tilde{t},\delta,\mu)=\dfrac{\mu\tilde{t}}{\delta}$ denoting the equilibrium profit outflow from each country to the tax haven.

The following proposition shows that contingent on the tax capacity and production technology, tax deductibility affects countries' tax revenues differently.

Proposition 3. *(i) A tax regime with lower deductibility can yield a larger tax revenue when tax capacity is sufficiently low. That is,* $\forall \mu_1, \mu_2$ *with* $0 \le \mu_1 < \mu_2 \le 1$, *there exists some* $\bar{\delta} > 0$ *such that* $\tilde{R}(\delta, \mu_1) > \tilde{R}(\delta, \mu_2)$ *holds when* $0 < \delta < \bar{\delta}$. *(ii) (The Diamond-Mirrlees production efficiency theorem) The equilibrium tax revenue is strictly increasing in deductibility parameter* μ *for all* $\mu \in [0,1]$ *when tax capacity is sufficiently high, marginal production cost is sufficiently low or output elasticity is sufficiently large.*

Proof. See Appendix A7.

For very low tax capacity, countries face severe profit shifting, and the revenue efficiency concern is strong. Part (i) shows that in this case limiting tax deductibility can be beneficial, although it causes production distortion. Intuitively, a lower deductibility reduces the tax sensitivity of profit shifting and increase the MNE's tax compliance. In Particular, as $\delta \to 0$, the effect of tax capacity on the equilibrium revenue converges to: $\lim_{\delta \to 0} \dfrac{\partial \tilde{R}(\delta, \mu)}{\partial \delta} = \dfrac{2[\bar{y} - \mu c(\bar{y})]^2}{9\mu^2}$ (cf. (A-27)), which strictly decreases with deductibility parameter μ. This implies that under very low tax capacity, a tax regime with broader tax base can tax the MNE more effectively, since the tax capacity exerts a larger effect on equilibrium tax revenue.

In contrast to part (i), part (ii) provides conditions under which the DM theorem is valid in the context of international tax competition and profit shifting. It implies that for high tax capacity, low marginal production cost or large output elasticity, the optimal tax policy is a pure profit tax that achieves full production efficiency, even though some profits go untaxed due to the MNE's profit shifting. Our result differs from Best et al. (2015, Proposition 1 and pp. 1322-1323), who argue that deviating from production efficiency is always desirable in the presence of tax evasion. To understand the rationale, given (31), we decompose the effect of tax deductibility on the equilibrium tax revenue as follows:

$$\frac{\partial \tilde{R}(\delta, \mu)}{\partial \mu} = \underbrace{\frac{\partial(\tilde{t}\,\tilde{\pi}^t(\tilde{t}, \delta, \mu))}{\partial \tilde{t}} \cdot \frac{\partial \tilde{t}(\delta, \mu)}{\partial \mu}}_{\text{indirect effect}} + \underbrace{\tilde{t} \cdot \frac{\partial \tilde{\pi}^t(\tilde{t}, \delta, \mu)}{\partial \mu}}_{\text{tax base effect}}.$$

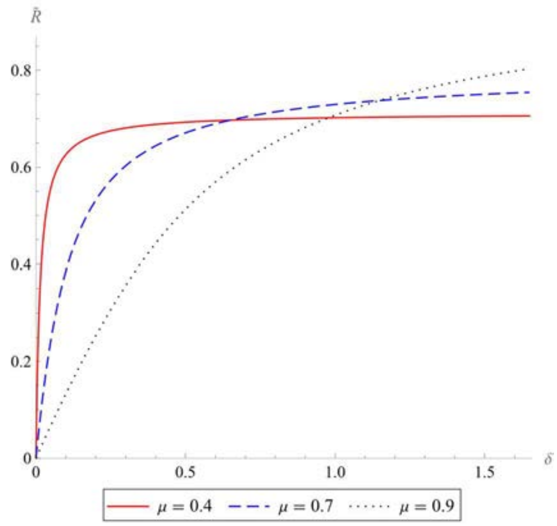
The indirect (tax rate) effect captures the change in tax revenue that is driven by the change of equilibrium tax rate. The tax base effect captures the change in tax revenue holding equilibrium tax rate constant. Due to the MNE's profit shifting, one country's corporate taxation creates a positive fiscal externality on the other country, i.e., $\dfrac{\partial(\tilde{t}\,\tilde{\pi}^t(\tilde{t}, \delta, \mu))}{\partial \tilde{t}} > 0$, $\forall \mu \in (0,1]$. Increasing tax deductibility has two opposite effects on the tax rate. Firstly, it raises the tax sensitivity of profit shifting, imposing a downward pressure on the tax rate. Secondly, it tends to reduce the EMTR and increase the output and taxable income in each country. This puts an *upward* pressure on the tax rate. Moreover, in response to a higher deductibility, the profit shifting potential is limited under high tax capacity, while the output increase is large under low marginal production cost or large output elasticity. In these cases, the positive effect of tax deductibility prevails such that the equilibrium tax rate strictly increases with the deductibility (cf. (A-32) and (A-33)).

On the other hand, the tax base effect can be further decomposed into three sub-effects:

$$\tilde{t}\cdot\frac{\partial\tilde{\pi}^t(\tilde{t},\delta,\mu)}{\partial\mu}=\tilde{t}\left[\underbrace{(1-\mu c'(\tilde{y}))\frac{\partial\tilde{y}(\tilde{t},\mu)}{\partial\mu}}_{production\ effect}-\underbrace{(c(\tilde{y})+\tilde{g})}_{deduction\ effect}-\underbrace{\mu\frac{\partial\tilde{g}(\tilde{t},\delta,\mu)}{\partial\mu}}_{shifting\ effect}\right].$$

The "production effect" is positive. A larger μ leads to a higher real output, which indicates a production efficiency gain and increases the tax base. The "deduction effect" is negative. A larger μ means that a larger share of costs is deductible, which narrows down the tax base. The "shifting effect" is also negative. When μ increases, more profits are shifted from each country to the tax haven, which indicates a revenue efficiency loss and reduces the tax base. Under high tax capacity, low marginal production cost or large output elasticity, the production effect can dominate the two negative effects such that a higher deductibility increases the equilibrium tax base (cf. (A-35)). Together with the positive indirect effect under these conditions, the DM theorem is extended to international profit shifting context.

(A) $\tilde{R}(\delta)$ with different μ



(B) $\tilde{R}(\mu)$ with different δ (C) $\tilde{R}(\mu)$ with different k

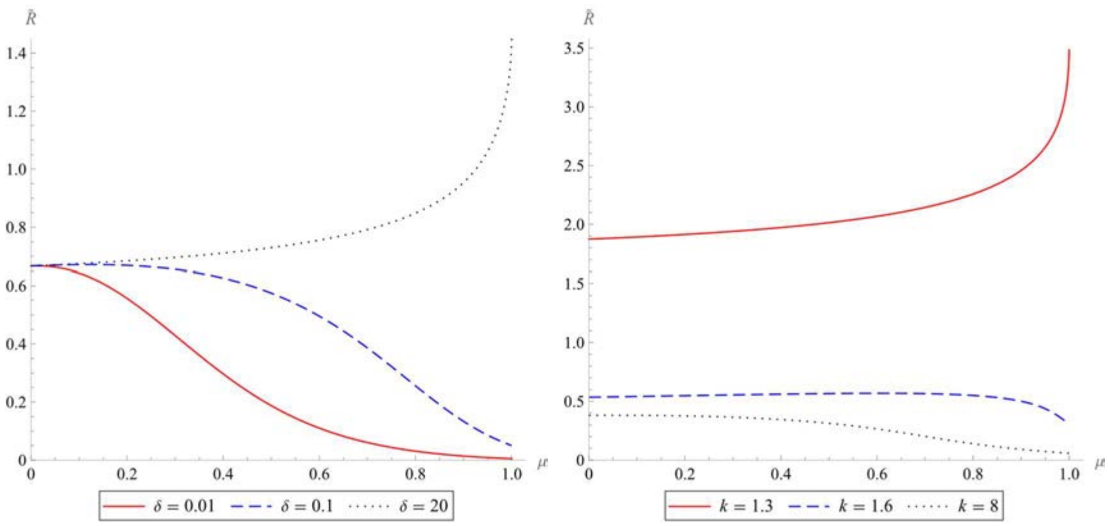


Fig. 2. The effects of tax capacity, tax deductibility and output elasticity on tax revenue.

Panel A illustrates equilibrium tax revenues under different deductibility rates and shows how tax rankings are related to the tax capacity. Panels B and C plot the relationship between equilibrium tax revenue and the deductibility under different tax capacities and different output elasticities, respectively. In Panels A and B, parameters are fixed as $\alpha = \frac{1}{6}$, $\beta = \frac{1}{2}$ and $k = \frac{3}{2}$. In Panel C, parameters are fixed as $\delta = 1$, $\alpha = \frac{1}{6}$ and $\beta = \frac{1}{2}$.

In the remainder of this section, we complement the analytical results with numerical simulations. Denote by μ^* the optimal deductibility that maximizes each country's equilibrium revenue, i.e., $\mu^* = \arg\max_{\mu \in [0,1]} \tilde{R}(\delta, \mu; \alpha, \beta, k)$. As Fig. 2 illustrates, the effects of the deductibility on equilibrium revenue are related to both tax capacity and production technology. Panel A compares three tax regimes with different tax deductibility. Under very low (high) tax capacity, the tax regime granting low (high) tax deductibility with $\mu = 0.4$ ($\mu = 0.9$) can yield larger revenue than the other two tax regimes. In Panel B, when tax capacity is not high, the optimal deductibility rate is interior, as shown by the solid curve ($\mu^* = 0.0132$) and the dashed curve ($\mu^* = 0.124$). The dotted curve indicates the case of sufficiently high tax capacity, where equilibrium tax revenue strictly increases with tax deductibility such that the optimal tax policy is a pure profit tax preserving full production efficiency. In Panel C, the solid curve shows the case of sufficiently high output elasticity, where the equilibrium revenue strictly increases with the deductibility. Otherwise, the optimal deductibility is interior and there is a bell-shaped relationship between equilibrium revenue and the deductibility, as illustrated by the dashed curve ($\mu^* = 0.608$) and dotted curve ($\mu^* = 0.0541$). It is also notable from Panels B and C that the optimal tax base is very close to the turnover tax base when the tax capacity or output elasticity is very low.

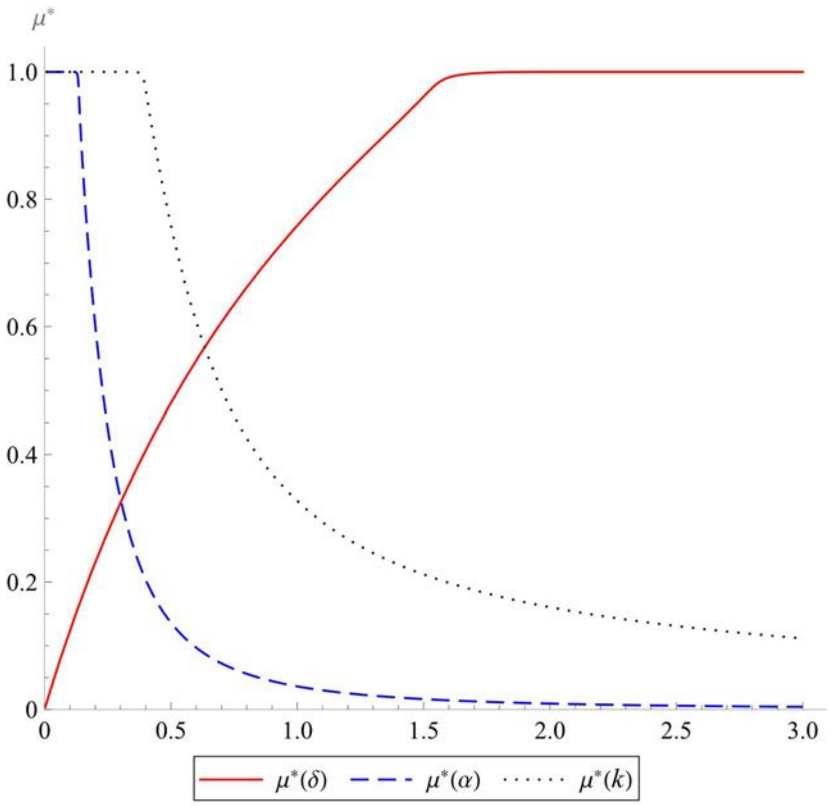


Fig. 3. The effects of δ, α and k on the optimal deductibility.

The parameter values are specified as: $\delta = 1$, $\alpha = \frac{1}{6}$, $\beta = \frac{1}{2}$ and $k = \frac{3}{2}$. To plot each curve, we fix the values of three parameters such that the optimal deductibility is a univariate function of the remaining parameter (e.g., $\mu^*(\delta)$ is plotted by setting $\alpha = \frac{1}{6}$, $\beta = \frac{1}{2}$ and $k = \frac{3}{2}$). For convenience, $\mu^*(k)$ is plotted in the $(k-1,\ \mu^*)$ space.

Fig. 3 illustrates how tax capacity and production technology affect the optimal tax deductibility. The optimal deductibility is increasing in δ, and a pure profit tax is optimal when δ is sufficiently large. By contrast, the optimal deductibility decreases with α and k. In particular, full deductibility is optimal when α is close to 0 or when k is close to 1. These results are in line with Panels B and C of Fig. 2. Intuitively, increasing tax capacity reduces the tax sensitivity of profit shifting and increases the revenue efficiency. Either a fall in the marginal production cost or a rise in the output elasticity indicates a larger production efficiency loss. In all these cases, the production efficiency concern becomes more important, making the optimal tax base closer to a pure profit tax base.

6. Conclusion

This paper has investigated the impact of tax deductibility on MNEs' behavior and countries' corporate taxes in the context of international profit shifting. We set up a formal tax competition model, in which two countries set corporate taxes to maximize their revenues in the presence of a tax haven levying zero tax rate. The MNE chooses output in each country and intra-firm transfer prices to maximize total after-tax profits.

We start from the polar cases in which costs in the affiliates of two countries are either fully tax deductible (a pure profit tax) or fully non-deductible (a turnover tax). Turnover tax can eliminate profit shifting incentives at the expense of production distortion. In contrast, profit tax (with either separate accounting or formula apportionment) can preserve production efficiency at the expense of tax base erosion. Under profit tax with formula apportionment, we identify an additional formula manipulation effect, which negatively affects countries' taxes. We show that the tax revenue comparisons of the three tax regimes depend on the MNE's production technology and countries' tax capacity. Specifically, profit tax with formula apportionment dominates the other two tax regimes only when the output elasticity is high and tax capacity is intermediate. In other cases, turnover tax (profit tax with separate accounting) can dominate under low (high) tax capacity. We then go beyond profit and turnover taxation to analyze the general case of tax deductibility and revisit the Diamond-Mirrlees production efficiency theorem in the presence of profit shifting. We show that a lower deductibility can yield larger tax revenue under very low tax capacity. However, it is also shown that equilibrium revenue strictly increases with tax deductibility when tax capacity is sufficiently high, marginal production cost is sufficiently low or output elasticity is sufficiently large. Simulations further indicate that the optimal tax deductibility is increasing in tax capacity and output elasticity, and decreasing in marginal production cost.

Our findings reveal that it is crucial for countries to maintain a balance between curbing profit shifting and improving production efficiency in order to tax multinationals effectively. The international tax design may differ in different industries, depending on countries' tax capacity and MNEs' production technologies. While it could be beneficial for African countries to shift from profit-based tax to turnover tax in the mining sector, it is not optimal to impose revenue-based DSTs on highly digitalized businesses. Due to the low marginal production cost and high output elasticity in the digital economy, the gain in revenue efficiency under turnover taxation is likely to be offset by the great loss of production efficiency.

Appendix A. Proofs.

A1. Proof of Lemma 2

From (12), the equilibrium profit outflow from each FA country to the tax haven is given by:

$$\hat{g} = \frac{\hat{t}}{\delta}. \tag{A-1}$$

Inserting (A-1) into (17) yields:

$$\bar{\pi} - \frac{3}{2}\hat{g} - \frac{\delta\hat{g}(\bar{\pi} - \hat{g})^2}{2(1-\delta\hat{g})\bar{y}^2 c''(\bar{y})} = 0, \tag{A-2}$$

from which we can obtain the following equation that implicitly determines $\hat{g}$:

$$\delta = \frac{\left(2\bar{\pi} - 3\hat{g}\right)\bar{y}^2 c''(\bar{y})}{\hat{g}\left[\left(\bar{\pi} - \hat{g}\right)^2 + \left(2\bar{\pi} - 3\hat{g}\right)\bar{y}^2 c''(\bar{y})\right]}. \tag{A-3}$$

Define $\xi(g) := \dfrac{\left(2\bar{\pi} - 3g\right)\bar{y}^2 c''(\bar{y})}{g\left[\left(\bar{\pi} - g\right)^2 + \left(2\bar{\pi} - 3g\right)\bar{y}^2 c''(\bar{y})\right]}$, $g \in (0, \dfrac{2\bar{\pi}}{3})$. For any given $\delta > 0$, notice that $\xi(2\bar{\pi}/3) = 0 < \delta < \lim_{g \to 0^+} \xi(g) = +\infty$ and that function $\xi(g)$ is continuous over interval $(0, \dfrac{2\bar{\pi}}{3})$. Then by the intermediate value theorem, there must exist some $\hat{g} \in (0, \dfrac{2\bar{\pi}}{3})$ such that $\xi(\hat{g}) = \delta$.

On the other hand, differentiating $\xi(g)$ w.r.t. g leads to:

$$\xi'(g) = -\frac{\left[2\left(\bar{\pi} - g\right)\left(\bar{\pi}^2 - 3g\bar{\pi} + 3g^2\right) + \left(2\bar{\pi} - 3g\right)^2 \bar{y}^2 c''(\bar{y})\right]\bar{y}^2 c''(\bar{y})}{g^2\left[\left(\bar{\pi} - g\right)^2 + \left(2\bar{\pi} - 3g\right)\bar{y}^2 c''(\bar{y})\right]^2} < 0. \tag{A-4}$$

(A-4) shows that $\xi(g)$ is strictly decreasing in g over $(0, \dfrac{2\bar{\pi}}{3})$, which guarantees a unique $\hat{g} \in (0, \dfrac{2\bar{\pi}}{3})$ satisfying $\xi(\hat{g}) = \delta$. In addition, it follows from (A-1) and (A-2) that: $\hat{t} \in (0,1) \Leftrightarrow \hat{g} \in (0, \dfrac{1}{\delta}) \Leftrightarrow \hat{g} \in (0, \dfrac{2\bar{\pi}}{3})$. Thus, for any $\delta > 0$, there exists a unique symmetric Nash equilibrium tax $\hat{t} \in (0,1)$ that is determined by (17).

Now we prove (ii). Given (A-4), it is obvious that:

$$\frac{d\hat{g}}{d\delta} = \frac{1}{\xi'(g)} < 0. \tag{A-5}$$

Hence, the equilibrium profit outflow from each FA country is strictly decreasing in the tax capacity.

Using (A-1) and (A-3), we rewrite the equilibrium tax $\hat{t}$ as a function of $\hat{g}$:

$$\hat{t} = \delta \hat{g} = \frac{\left(2\bar{\pi} - 3\hat{g}\right)\bar{y}^2 c''(\bar{y})}{\left(\bar{\pi} - \hat{g}\right)^2 + \left(2\bar{\pi} - 3\hat{g}\right)\bar{y}^2 c''(\bar{y})}, \tag{A-6}$$

from which we can derive:

$$\frac{d\hat{t}}{d\delta} = \frac{d\hat{t}}{d\hat{g}} \cdot \frac{d\hat{g}}{d\delta} = \frac{\left(\bar{\pi} - \hat{g}\right)\left(\bar{\pi} - 3\hat{g}\right)\bar{y}^2 c''(\bar{y})}{\left[\left(\bar{\pi} - \hat{g}\right)^2 + \left(2\bar{\pi} - 3\hat{g}\right) y^2 c''(\bar{y})\right]^2} \cdot \frac{d\hat{g}}{d\delta}. \tag{A-7}$$

Then it follows from (A-4), (A-5) and (A-7) that:

$$\frac{d\hat{t}}{d\delta} \gtreqless 0 \iff \hat{g} \gtreqless \frac{\bar{\pi}}{3} \iff \delta \lesseqgtr \delta^*, \tag{A-8}$$

where $\delta^* := \xi(\frac{\bar{\pi}}{3}) = \frac{27k}{(4+9k)\bar{\pi}}$.

Lastly, we prove (iii). Notice that the equilibrium revenue of each FA country can be expressed as: $\hat{R} = \hat{t}(\bar{\pi} - \hat{g})$. Given (A-5) – (A-7), we can derive the following:
$$\frac{d\hat{R}}{d\delta} = \hat{t}(-\frac{d\hat{g}}{d\delta}) + \left(\bar{\pi} - \hat{g}\right)\frac{d\hat{t}}{d\delta} = -\frac{\left[\left(\bar{\pi} - \hat{g}\right)^2 \bar{\pi} + \left(2\bar{\pi} - 3\hat{g}\right)^2 \bar{y}^2 c''(\bar{y})\right]\bar{y}^2 c''(\bar{y})}{\left[\left(\bar{\pi} - \hat{g}\right)^2 + \left(2\bar{\pi} - 3\hat{g}\right)\bar{y}^2 c''(\bar{y})\right]^2} \cdot \frac{d\hat{g}}{d\delta} > 0 \quad .$$

Q.E.D.

A2. Proof of Proposition 1

Notice that tax capacity thresholds δ^{SA} and δ^{FA} are determined by equations $\tilde{R}(\delta^{SA}) = R^T$ and $\hat{R}(\delta^{FA}) = R^T$, respectively. We first derive the expression for δ^{SA}. Using (8), (23) and Lemma 1(iii), we have the following:

$$\delta^{SA} \lesseqgtr \frac{3}{\bar{\pi}} \iff \tilde{R}(\delta^{SA}) \lesseqgtr \tilde{R}(\frac{3}{\bar{\pi}}) \iff R^T \lesseqgtr \tilde{R}(\frac{3}{\bar{\pi}}) \iff k^{\frac{1}{1-k}}\bar{\pi} \lesseqgtr \frac{2\bar{\pi}}{3} \iff k \lesseqgtr 4.939, \tag{A-9}$$

where 4.939 is the approximate solution to $k^{\frac{1}{1-k}} - \frac{2}{3} = 0$.

Then it follows from (8), (23) and (A-9) that:

$$\tilde{R}(\delta^{SA}) = R^T \quad \Leftrightarrow \quad \begin{cases} \dfrac{2\overline{\pi}^2\delta^{SA}}{9} = k^{\frac{1}{1-k}}\overline{\pi} & \text{if } 1 < k < 4.939 \\ \overline{\pi} - \dfrac{1}{\delta^{SA}} = k^{\frac{1}{1-k}}\overline{\pi} & \text{if } k \geq 4.939 \end{cases} \quad \Leftrightarrow$$

$$\delta^{SA} = \begin{cases} \dfrac{9k^{\frac{1}{1-k}}}{2\overline{\pi}} & \text{if } 1 < k < 4.939 \\ \dfrac{1}{\overline{\pi}(1-k^{\frac{1}{1-k}})} & \text{if } k \geq 4.939 \end{cases}. \qquad \text{(A-10)}$$

We now derive δ^{FA} in closed form. Denote by g^{FA} the equilibrium profit outflow in each country under tax capacity level δ^{FA}. Using (A-3), (23) and $\overline{y}^2c''(\overline{y}) = k\overline{\pi}$, we get the following:

$$\hat{R}(\delta^{FA}) = R^T \quad \Leftrightarrow \quad \frac{k\overline{\pi}(\overline{\pi} - g^{FA})\left(2\overline{\pi} - 3g^{FA}\right)}{\left(\overline{\pi} - g^{FA}\right)^2 + k\overline{\pi}\left(2\overline{\pi} - 3g^{FA}\right)} = k^{\frac{1}{1-k}}\overline{\pi} \quad \Leftrightarrow$$

$$f(g^{FA}) := (g^{FA})^2\left(3k^{\frac{k}{k-1}} - 1\right) + g^{FA}\left(2 + 3k - 5k^{\frac{k}{-1+k}}\right)\overline{\pi} + \left(2k^{\frac{k}{-1+k}} - 2k - 1\right)\overline{\pi}^2 = 0 \quad \Leftrightarrow$$

$$g^{FA} = \frac{\left(5k^{\frac{k}{k-1}} - 3k - 2 - \sqrt{\Omega}\right)\overline{\pi}}{6k^{\frac{k}{k-1}} - 2}, \text{ where } \Omega := k^{\frac{2k}{k-1}} - 6k^{\frac{2k-1}{k-1}} + k(4+9k) > 0, \ \forall k > 1.$$

Plugging g^{FA} into (A-3) yields the expression for δ^{FA}:

$$\delta^{FA} = \frac{2k^{\frac{1}{1-k}}\left(3k^{\frac{k}{k-1}} - 1\right)^2\left(9k - 3k^{\frac{k}{k-1}} + 3\sqrt{\Omega} + 2\right)}{\overline{\pi}\left(5k^{\frac{k}{k-1}} - 3k - \sqrt{\Omega} - 2\right)\left(9k\left(3k+1\right) - \left(9k-1\right)k^{\frac{k}{k-1}} + (9k+1)\sqrt{\Omega}\right)}. \qquad \text{(A-11)}$$

For a given tax capacity level δ, it directly follows from Lemma 1(iii) and Lemma 2(iii) that:

$$\tilde{R}(\delta) \lesseqqgtr R^T \quad \Leftrightarrow \quad \delta \lesseqqgtr \delta^{SA}, \text{ and } \hat{R}(\delta) \lesseqqgtr R^T \quad \Leftrightarrow \quad \delta \lesseqqgtr \delta^{FA}. \qquad \text{(A-12)}$$

Notably, when $\delta = +\infty$, it follows from (11) and (12) that $\forall(t_1, t_2) \in [0,1]^2$, $g_{ij} \equiv 0$ for $1 \leq i < j \leq 3$. This corresponds to the case where the tax haven is absent. So, in an economy without tax haven each country's equilibrium revenue under FA is $\hat{R}(+\infty)$. Obviously, $\hat{R}(+\infty) > R^T$ always holds since $\delta = +\infty > \delta^{FA}$.

Now we prove (ii). Recall that $\bar{\pi}$ is strictly decreasing in α and k, and that $k^{\frac{1}{1-k}}$ is strictly increasing in k. These properties together with (A-10) lead to the effects of α and k on threshold δ^{SA}.

On the other hand, define

$$\phi(k) := \frac{2k^{\frac{1}{1-k}}\left(3k^{\frac{k}{k-1}}-1\right)^2\left(9k-3k^{\frac{k}{k-1}}+3\sqrt{\Omega}+2\right)}{\left(5k^{\frac{k}{k-1}}-3k-\sqrt{\Omega}-2\right)\left(9k\left(3k+1\right)-\left(9k-1\right)k^{\frac{k}{k-1}}+(9k+1)\sqrt{\Omega}\right)}, \quad k>1.$$

By numerical method, we can graph in the $(1/k, \phi'(k))$ space function $\phi'(k)$ with $k \in (1, +\infty)$:

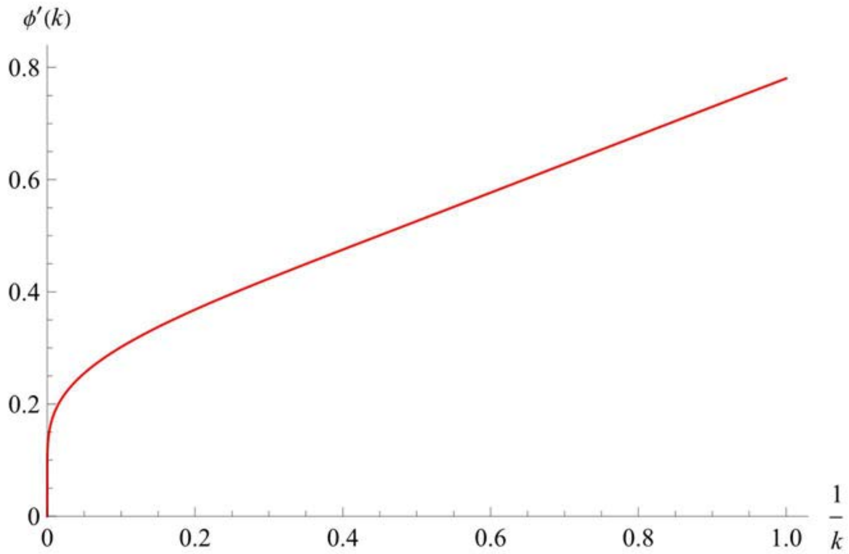


Fig. A1. The graph of $\phi'(k)$.

Fig. A1 illustrates that $\phi(k)$ is strictly increasing in k for all $k>1$. Besides, $\bar{\pi}$ is strictly decreasing in α and in k. Given these properties and (A-11), the comparative static effects on δ^{FA} immediately follow. **Q.E.D.**

A3. Proof of Lemma 4

We analyze two cases: (i) $0<\delta<\frac{3}{\bar{\pi}}$, and (ii) $\delta \geq \frac{3}{\bar{\pi}}$, separately.

Case (i): $0<\delta<\frac{3}{\bar{\pi}}$. Recall that $\hat{g} \in (0, \frac{2\bar{\pi}}{3})$. Then using (7), (8), (A-1) and (A-8), we can derive the following:

$$\tilde{R} \lesseqqgtr \hat{R} \iff \frac{2\bar{\pi}^2\delta}{9} \lesseqqgtr \delta\hat{g}(\bar{\pi}-\hat{g}) \iff \hat{g} \gtreqqless \frac{\bar{\pi}}{3} \iff \hat{t} \gtreqqless \tilde{t} \iff \delta \lesseqqgtr \delta^*, \tag{A-13}$$

where $\delta^* := \dfrac{27k}{(4+9k)\bar{\pi}} < \dfrac{3}{\bar{\pi}}$.

Case (ii): $\delta \geq \dfrac{3}{\bar{\pi}}$. It follows from (7) and Lemma 2(i) that:

$$\tilde{t} = 1 > \hat{t} > 0. \tag{A-14}$$

Notice that for a given value of δ, $t(\bar{\pi} - \dfrac{t}{\delta})$ is strictly increasing in t over interval $[0, \delta\bar{\pi}/2]$. Besides, $[0,1] \subseteq [0, \delta\bar{\pi}/2]$ since $\delta\bar{\pi} \geq 3$. Thus, we obtain:

$$\forall \delta \geq \frac{3}{\bar{\pi}},\ \tilde{R} = \tilde{t}(\bar{\pi} - \frac{\tilde{t}}{\delta}) > \hat{t}(\bar{\pi} - \frac{\hat{t}}{\delta}) = \hat{R}. \tag{A-15}$$

Combining (A-13) – (A-15) leads to Lemma 4. **Q.E.D.**

A4. Proof of Proposition 2

We first rank δ^{SA}, δ^{FA} and δ^*, conditional on the value of k. Plugging $\delta = \delta^*$ into (8) and using (23) leads to:

$$\tilde{R}(\delta^*) \gtreqless R^T \ \Leftrightarrow\ \frac{6k\bar{\pi}}{4+9k} \gtreqless k^{\frac{1}{1-k}}\bar{\pi} \ \Leftrightarrow\ k \lesseqgtr 3.089, \tag{A-16}$$

where 3.089 is the approximate solution to $4k^{\frac{k}{1-k}} + 9k^{\frac{1}{1-k}} - 6 = 0$.

On the other hand, using (A-12) and Lemma 4, we can derive:

$$\tilde{R}(\delta^*) \gtreqless R^T \ \Rightarrow\ \delta^* \gtreqless \delta^{SA} \ \Rightarrow\ \hat{R}(\delta^{SA}) \gtreqless \tilde{R}(\delta^{SA}) = R^T \ \Rightarrow\ \delta^{SA} \gtreqless \delta^{FA}. \tag{A-17}$$

Combining (A-16) and (A-17) yields the ranking of the three thresholds:

$$\delta^* \gtreqless \delta^{SA} \gtreqless \delta^{FA} \ \Leftrightarrow\ k \lesseqgtr 3.089. \tag{A-18}$$

Given Proposition 1(i), Lemma 4 and (A-18), part (i) and (ii) immediately follow.

Now we prove (iii). Given the technological parameters α and k, denote by $A(\alpha,k)$ ($B(\alpha,k)$) the set of tax capacity in which SA (turnover tax) is superior to the other two tax regimes. Then it directly follows from Proposition 2(i) that:

$$A(\alpha,k) = \left\{\delta : \delta > \delta^*(\alpha,k) \text{ and } 1 < k < 3.089\right\} \cup \left\{\delta : \delta > \delta^{SA}(\alpha,k) \text{ and } k \geq 3.089\right\},$$

$$B(\alpha,k) = \left\{\delta : \delta < \delta^{FA}(\alpha,k) \text{ and } 1 < k < 3.089\right\} \cup \left\{\delta : \delta < \delta^{SA}(\alpha,k) \text{ and } k \geq 3.089\right\}.$$

Recall from Lemma 2(ii) and Proposition 1(ii) that $\delta^{SA}(\alpha,k)$, $\delta^{FA}(\alpha,k)$ and $\delta^*(\alpha,k)$ are strictly increasing in α and k. Using these properties and (A-18), it is straightforward to show the following:

$\forall \alpha_1 > \alpha_2 > 0$, $A(\alpha_1,k) \subset A(\alpha_2,k)$ and $B(\alpha_1,k) \supset B(\alpha_2,k)$;

$\forall k_1 > k_2 > 1$, $A(\alpha,k_1) \subset A(\alpha,k_2)$ and $B(\alpha,k_1) \supset B(\alpha,k_2)$.

Therefore, the scope within which SA (turnover tax) dominates the other two tax regimes strictly decreases (increases) with α and k. **Q.E.D.**

A5. The existence and uniqueness of symmetric equilibrium under SA

As two polar cases (i.e., $\mu \in \{0,1\}$) have been studied in Section 3, we here focus on the case of partial deductibility with $\mu \in (0,1)$. Given the specific functional form of $c(.)$, it follows from (26) that $\tilde{t} \le \dfrac{1-\beta}{1-\beta\mu} \le 1$. Define $F(t,\mu,\delta) := y - \mu c(y) - \dfrac{t(1-\mu)^2}{(1-\mu t)^3 c''(y)} - \dfrac{3\mu^2 t}{\delta}$, where y is determined by (26). It is obvious that: $\forall \mu \in (0,1)$, $F(0,\mu,\delta) = \bar{y} - \mu c(\bar{y}) > 0$ and $\lim_{t \to ((1-\beta)/(1-\beta\mu))^-} F(t,\mu,\delta) < 0$. This property together with the continuity of $F(t,\mu,\delta)$ in t ensures the existence of a Nash equilibrium with tax rate $\tilde{t} \in \left(0, \dfrac{1-\beta}{1-\beta\mu}\right)$.

On the other hand, solving for $\tilde{y}$ from (26) yields $\tilde{y} = \left[\dfrac{1-\tilde{t}-\beta(1-\mu\tilde{t})}{\alpha(1-\mu\tilde{t})}\right]^{\frac{1}{k-1}}$. Then plugging this into (30) and rearranging terms gives:

$$H(k,\beta,\mu,\tilde{t}) = \frac{3\mu^2\tilde{t}}{\delta} \cdot \alpha^{\frac{1}{k-1}} \left(\frac{1-\tilde{t}}{1-\mu\tilde{t}} - \beta\right)^{\frac{k-2}{k-1}} > 0, \tag{A-19}$$

where $H(k,\beta,\mu,\tilde{t}) := (1-\beta\mu)\left(\dfrac{1-\tilde{t}}{1-\mu\tilde{t}} - \beta\right) - \dfrac{\mu}{k}\left(\dfrac{1-\tilde{t}}{1-\mu\tilde{t}} - \beta\right)^2 - \dfrac{\tilde{t}(1-\mu)^2}{(k-1)(1-\mu\tilde{t})^3}$ is strictly increasing in k.

Given (26), (28) and the specification of $c(.)$, computing $\dfrac{\partial F(t,\mu,\delta)}{\partial t}$ at Nash equilibrium tax rate $\tilde{t}$ and simplifying yields:

$$\frac{\partial F(\tilde{t},\mu,\delta)}{\partial \tilde{t}} = -\frac{(1-\mu)^2 \Phi_1}{(k-1)(1-\mu\tilde{t})^4[1-\tilde{t}-\beta(1-\mu\tilde{t})]c''(\tilde{y})} - \frac{3\mu^2}{\delta}, \tag{A-20}$$

where $\Phi_1 := (k-1)(2+\mu\tilde{t})\left[1-\tilde{t}-\beta(1-\mu\tilde{t})\right] + \tilde{t}(1-\mu)(k-2)$.

Now we show that it cannot be the case that $\Phi_1 \le 0$. If it were true, we would obtain:

$$k \le \frac{(2+\mu\tilde{t})\left[1-\tilde{t}-\beta(1-\mu\tilde{t})\right] + 2\tilde{t}(1-\mu)}{(2+\mu\tilde{t})\left[1-\tilde{t}-\beta(1-\mu\tilde{t})\right] + \tilde{t}(1-\mu)} =: \underline{k} \quad \Rightarrow$$

$$0 < H(k,\beta,\mu,\tilde{t}) \le H(\underline{k},\beta,\mu,\tilde{t}) = \frac{-(1-\mu)\Phi_2}{(1-\mu\tilde{t})^3\left[(2+\mu\tilde{t})(1-\tilde{t}-\beta(1-\mu\tilde{t})) + 2\tilde{t}(1-\mu)\right]} < 0,$$

where

$$\Phi_2 := (1-\mu\tilde{t})^2\left(3\mu^2\tilde{t}^2+4\mu\tilde{t}+2\right)\beta^2+(1-\mu\tilde{t})\left(6\mu^2\tilde{t}^3-\mu^2\tilde{t}^2+3\mu\tilde{t}^2-4\mu\tilde{t}-4\right)\beta$$
$$+(3\mu^2\tilde{t}^4-\mu^2\tilde{t}^3-\mu\tilde{t}^3-3\mu\tilde{t}^2+2)>0.$$

From the above contradiction, it must be that $\Phi > 0$. Then it follows from (A-20) that:

$$\forall\mu\in(0,1),\ \frac{\partial F(\tilde{t},\mu,\delta)}{\partial\tilde{t}}<0. \tag{A-21}$$

(A-21) states that the partial derivative of $-F(t,\mu,\delta)$ w.r.t. t is positive at *all candidate equilibria* satisfying $F(\tilde{t},\mu,\delta)=0$. Then by the Poincaré-Hopf index theorem (see Vives, 2000), the uniqueness of $\tilde{t}$ directly follows.

A6. The range of the equilibrium output under SA

In Section 3, we have obtained each country's equilibrium output under turnover and profit taxation, i.e., $\tilde{y}=\underline{y}$ if $\mu=0$, and $\tilde{y}=\overline{y}$ if $\mu=1$. Here we investigate the partial deductibility with $\mu\in(0,1)$. Given (26) and $\tilde{t}\in\left(0,\frac{1-\beta}{1-\beta\mu}\right)$, it is obvious that: $\forall\mu\in(0,1)$, $\beta<c'(\tilde{y})<1$ and $\tilde{y}<\overline{y}$. On the other hand, it follows from (26) that $\tilde{t}=\frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}$. Inserting this expression into (A-19) yields:

$$H\left(k,\beta,\mu,\frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}\right)=(1-\beta\mu)\left[c'(\tilde{y})-\beta\right]-\frac{\mu}{k}\left[c'(\tilde{y})-\beta\right]^2-\frac{\left[1-c'(\tilde{y})\right]\left[1-\mu c'(\tilde{y})\right]^2}{(k-1)(1-\mu)}>0$$

Now we prove by contradiction that $\forall\mu\in(0,1)$, $\tilde{y}>\underline{y}$. Recall that $H\left(k,\beta,\mu,\frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}\right)$ is strictly increasing in k. If there were some $\mu\in(0,1)$ such that $\tilde{y}\le\underline{y}$, then we would derive the following:

$$\tilde{y}\le\underline{y}\ \Rightarrow\ c'(\tilde{y})\le c'(\underline{y})=\frac{1-\beta}{k}+\beta\ \Rightarrow\ k\le\frac{1-\beta}{c'(\tilde{y})-\beta}\ \Rightarrow$$

$$0<H\left(k,\beta,\mu,\frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}\right)\le H\left(\frac{1-\beta}{c'(\tilde{y})-\beta},\beta,\mu,\frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}\right)=-\frac{\mu(1-\mu\beta)\left[c'(\tilde{y})-\beta\right]\left[1-c'(\tilde{y})\right]^2}{(1-\mu)(1-\beta)}<0,$$

which is a contradiction. Therefore, we must have:

$$\underline{y}<\tilde{y}<\overline{y},\ \forall\mu\in(0,1). \tag{A-22}$$

A7. Proof of Proposition 3

Given (26), (28), (30) and (31), the effect of tax capacity on each country's equilibrium revenue is:

$$\frac{\partial\tilde{R}(\delta,\mu)}{\partial\delta}=\frac{\partial(\tilde{t}\,\tilde{\pi}^t(\tilde{t},\delta,\mu))}{\partial\tilde{t}}\cdot\frac{\partial\tilde{t}(\delta,\mu)}{\partial\delta}+\tilde{t}\cdot\frac{\partial\tilde{\pi}^t(\tilde{t},\delta,\mu)}{\partial\delta}, \tag{A-23}$$

where

$$\frac{\partial(\tilde{t}\,\tilde{\pi}^t(\tilde{t},\delta,\mu))}{\partial\tilde{t}}=\tilde{y}-\mu c(\tilde{y})-\frac{\mu^2\tilde{t}}{\delta}+\tilde{t}\left[(1-\mu c'(\tilde{y}))\frac{\partial\tilde{y}(\tilde{t},\mu)}{\partial\tilde{t}}-\frac{\mu^2}{\delta}\right]=\frac{\mu^2\tilde{t}}{\delta}>0, \tag{A-24}$$

$$\frac{\partial\tilde{\pi}^{t}(\tilde{t},\delta,\mu)}{\partial\delta}=\frac{\mu^{2}\tilde{t}}{\delta^{2}}>0\,. \tag{A-25}$$

On the other hand, using (26) and (30), we can derive the following:

$$\lim_{\delta\to0}\tilde{t}(\delta,\mu)=0\ ,\ \lim_{\delta\to0}\tilde{y}(\tilde{t},\mu)=\overline{y}\ ,\ \text{and}\ \lim_{\delta\to0}\frac{\tilde{t}(\delta,\mu)}{\delta}=\lim_{\delta\to0}\frac{\partial\tilde{t}(\delta,\mu)}{\partial\delta}=\frac{\overline{y}-\mu c(\overline{y})}{3\mu^{2}}\ ,\ \forall\mu\in(0,1]\,, \tag{A-26}$$

where the first equality in the third part of (A-26) holds by L'Hôpital's rule.

Taking the limit on both sides of (A-23) as $\delta\to0$ and using (A-24) – (A-26) yields:

$$\lim_{\delta\to0}\frac{\partial\tilde{R}(\delta,\mu)}{\partial\delta}=\lim_{\delta\to0}\left(\frac{\mu^{2}\tilde{t}}{\delta}\cdot\frac{\partial\tilde{t}(\delta,\mu)}{\partial\delta}+\frac{\mu^{2}\tilde{t}^{2}}{\delta^{2}}\right)=\frac{2[\overline{y}-\mu c(\overline{y})]^{2}}{9\mu^{2}},\ \ \forall\mu\in(0,1]\,. \tag{A-27}$$

(A-27) indicates that $\lim_{\delta\to0}\frac{\partial\tilde{R}(\delta,\mu)}{\partial\delta}$ is strictly decreasing in μ, which leads to the following:

$\forall\mu_1,\mu_2\in(0,1]$ with $\mu_1<\mu_2$, there exists some $\overline{\delta}>0$ such that $\frac{\partial}{\partial\delta}[\tilde{R}(\delta,\mu_1)-\tilde{R}(\delta,\mu_2)]>0$ when $0<\delta<\overline{\delta}$ $\Rightarrow$

$\tilde{R}(\delta,\mu_1)-\tilde{R}(\delta,\mu_2)>\lim_{\delta\to0}[\tilde{R}(\delta,\mu_1)-\tilde{R}(\delta,\mu_2)]=0$ when $0<\delta<\overline{\delta}$.

For the special case where $\mu_1=0<\mu_2\le1$, we have $\tilde{R}(\delta,0)=R^{T}>0=\lim_{\delta\to0}\tilde{R}(\delta,\mu_2)$, where R^{T} is independent of tax capacity $\Rightarrow$ there exists some $\overline{\delta}>0$ such that $\tilde{R}(\delta,0)>\tilde{R}(\delta,\mu_2)$ when $0<\delta<\overline{\delta}$.

Now we prove (ii). For simplicity, in what follows we assume that $\beta=0$, while our result still holds for $0<\beta<1$. Given (31), the effect of the deductibility μ on each country's equilibrium tax revenue under SA is:

$$\frac{\partial\tilde{R}(\delta,\mu)}{\partial\mu}=\underbrace{\frac{\partial(\tilde{t}\,\tilde{\pi}^{t}(\tilde{t},\delta,\mu))}{\partial\tilde{t}}\cdot\frac{\partial\tilde{t}(\delta,\mu)}{\partial\mu}}_{\text{indirect effect}}+\underbrace{\tilde{t}\cdot\frac{\partial\tilde{\pi}^{t}(\tilde{t},\delta,\mu)}{\partial\mu}}_{\text{tax base effect}}\,. \tag{A-28}$$

We focus on the partial deductibility with $\mu\in(0,1)$, and proceed in two steps to show that both the indirect effect and tax base effect are positive under high tax capacity, low marginal production cost or large output elasticity.

Step 1:

Given (29), (26), (30) and the production cost function specification, differentiating $F(\tilde{t},\mu,\delta)$ w.r.t. μ and then heavily rearranging terms gives:

$$\frac{\partial F(\tilde{t},\mu,\delta)}{\partial\mu}=\frac{(1-\mu\tilde{t})}{1-\mu}\left[\tilde{y}c'(\tilde{y})-c(\tilde{y})-\frac{3\mu\tilde{t}(2-\mu-\mu\tilde{t})}{(1-\mu\tilde{t})\delta}\right]+\frac{\tilde{t}\,(1-\mu)}{(1-\mu\tilde{t})^{4}(k-1)c''(\tilde{y})}\left[2k(1-\tilde{t})+\tilde{t}(\mu+1)-2\right]. \tag{A-29}$$

Notice that $\tilde{y}c'(\tilde{y})-c(\tilde{y})$ is strictly increasing in $\tilde{y}$, and recall from (A-22) that $\tilde{y}>\underline{y}$. Besides, $\frac{2-\mu-\mu\tilde{t}}{1-\mu\tilde{t}}=1+\frac{1-\mu}{1-\mu\tilde{t}}\in(1,2)$. Then regarding the first term on the RHS of (A-29), we have:

$$\tilde{y}c'(\tilde{y})-c(\tilde{y})-\frac{3\mu\tilde{t}(2-\mu-\mu\tilde{t})}{(1-\mu\tilde{t})\delta}>0\ \ \text{if}\ \ \delta\ge\frac{6}{\underline{y}c'(\underline{y})-c(\underline{y})}\,. \tag{A-30}$$

Now we argue that it cannot be the case that $2k(1-\tilde{t})+\tilde{t}(\mu+1)-2\leq 0$. If it were true, then we would have: $k\leq\dfrac{2-\tilde{t}(\mu+1)}{2(1-\tilde{t})}$ $\Rightarrow$

$0<H(k,0,\mu,\tilde{t})\leq H\left(\dfrac{2-\tilde{t}(\mu+1)}{2(1-\tilde{t})},0,\mu,\tilde{t}\right)=-\dfrac{(1-\tilde{t})(1-\mu)\left(2-\tilde{t}-\mu^2\tilde{t}^3\right)}{(1-\mu\tilde{t})^3\left(2-\tilde{t}-\mu\tilde{t}\right)}<0$, which is a contradiction.

Therefore, given (A-29) and (A-30) we get:

$$\forall\mu\in(0,1),\ \frac{\partial F(t,\mu,\delta)}{\partial\mu}>0\ \text{ if }\ \delta\geq\frac{6}{\underline{y}c'(\underline{y})-c(\underline{y})}. \tag{A-31}$$

Given $F(\tilde{t},\mu,\delta)=0$, using (A-21), (A-31) and the implicit function theorem yields:

$$\frac{\partial\tilde{t}(\mu,\delta)}{\partial\mu}=-\frac{\partial F(\tilde{t},\mu,\delta)/\partial\mu}{\partial F(\tilde{t},\mu,\delta)/\partial\tilde{t}}>0\ \text{ if }\ \delta\geq\frac{6}{\underline{y}c'(\underline{y})-c(\underline{y})}. \tag{A-32}$$

Moreover, using (22), we get: $\underline{y}c'(\underline{y})-c(\underline{y})=(k-1)\alpha^{\frac{1}{1-k}}k^{\frac{2k-1}{1-k}}=:\varphi(\alpha,k)$. It is readily verified that $\varphi(\alpha,k)$ is strictly decreasing in α and k, with $\lim_{k\to1^+}\varphi(\alpha,k)=+\infty$ and $\lim_{k\to+\infty}\varphi(\alpha,k)=0$. Then we can derive that:

$$\delta\geq\frac{6}{\underline{y}c'(\underline{y})-c(\underline{y})}\ \Leftrightarrow\ \alpha\leq\left[\frac{\delta(k-1)}{6}\right]^{k-1}k^{1-2k}\ \Leftrightarrow\ k\leq\bar{k}, \tag{A-33}$$

where $\bar{k}$ is the unique value of k satisfying $\varphi(\alpha,k)=\dfrac{6}{\delta}$.

Given (A-24), (A-32) and (A-33), the indirect effect of tax deductibility is positive for a high tax capacity, low marginal production cost or large output elasticity.

Step 2:

Given (26) and (29) – (31), computing the tax base effect and then rearranging terms gives:

$$\begin{aligned}\tilde{t}\cdot\frac{\partial\tilde{\pi}^t(\tilde{t},\delta,\mu)}{\partial\mu}&=\tilde{t}\left[\underbrace{(1-\mu c'(\tilde{y}))\frac{\partial\tilde{y}(\tilde{t},\mu)}{\partial\mu}}_{\text{production effect}}-\underbrace{(c(\tilde{y})+\tilde{g})}_{\text{deduction effect}}-\underbrace{\mu\frac{\partial\tilde{g}(\tilde{t},\delta,\mu)}{\partial\mu}}_{\text{shifting effect}}\right]\\&=\frac{\tilde{t}(1-\mu\tilde{t})}{1-\mu}\left[\tilde{y}c'(\tilde{y})-c(\tilde{y})-\frac{\mu\tilde{t}(2+\mu-3\mu\tilde{t})}{(1-\mu\tilde{t})\delta}\right].\end{aligned} \tag{A-34}$$

Note that $\dfrac{2+\mu-3\mu\tilde{t}}{1-\mu\tilde{t}}=3-\dfrac{1-\mu}{1-\mu\tilde{t}}\in(2,3)$, and recall from (A-22) that $\tilde{y}>\underline{y}$. Then it immediately follows from (A-34) that:

$$\forall\mu\in(0,1),\ \tilde{t}\cdot\frac{\partial\tilde{\pi}^t(\tilde{t},\delta,\mu)}{\partial\mu}>0\ \text{ if }\ \delta\geq\frac{3}{\underline{y}c'(\underline{y})-c(\underline{y})}. \tag{A-35}$$

Combining (A-24), (A-28), (A-32) and (A-35) leads to:

$$\forall\mu\in(0,1),\ \frac{\partial\tilde{R}(\delta,\mu)}{\partial\mu}>0\ \text{ if }\ \delta\geq\frac{6}{\underline{y}c'(\underline{y})-c(\underline{y})}. \tag{A-36}$$

Given (A-33), (A-36) and the continuity of $\tilde{R}(\delta,\mu)$ at $\mu = 0,1$, Proposition 3(ii) directly follows. **Q.E.D.**

Appendix B. The impact of tax deductibility under FA

Suppose as in Subsection 3.2 that the two countries form a FA union, while the tax haven sticks to SA. Using (1) and (24) gives the MNE's consolidated tax base within the FA union:

$$\pi^c := \sum_{i=1}^{2}\left[y_i - \mu\left(c(y_i) + \sum_{1\le j\le 3, j\ne i} g_{ij} \right)\right] = \sum_{i=1}^{2}\left[y_i - \mu(c(y_i) + g_{i3})\right]. \tag{B-1}$$

The tax base of the MNE in country i is characterized by (10), where π^c is given by (B-1). Using (1), (2), (10) and (B-1), the MNE's total after-tax profit under FA amounts to:

$$\Pi^{FA} = (1-\tau)\pi^c - (1-\mu)\sum_{i=1}^{2} c(y_i) + \mu\sum_{i=1}^{2} g_{i3} - \sum_{1\le i<j\le 3}\frac{\delta}{2}g_{ij}^2,$$

where $\tau := \sum_{i=1}^{2}\lambda_i t_i$ is the effective tax rate within the FA union.

Solving the first-order conditions of the MNE's profit maximization yields:

$$g_{12} = g_{21} = 0, \tag{B-2}$$

$$g_{i3} = \frac{\mu\tau}{\delta}, \quad i = 1,2, \tag{B-3}$$

$$c'(y_i) = \frac{1-\tau}{1-\mu\tau} + \frac{(\tau - t_i)\pi^c}{(1-\mu\tau)\sum_{i=1}^{2} y_i}, \quad i = 1,2. \tag{B-4}$$

As shown by (B-2), the statement that profit shifting within the FA union can be eliminated still holds. On the other hand, (B-3) indicates that the profit shifted from each country to the tax haven falls with a broader tax base. In particular, profit shifting can be completely removed under a turnover tax base (i.e., $\mu = 0$). In the symmetric situation with equal tax rates $t_i = t = \tau$ ($i = 1,2$), it follows from (B-4) that the real output is equal in the two countries, which is still determined by (26) (with country subscripts dropped due to symmetry).

Totally differentiating (B-3) and (B-4), together with the symmetry property, yields the comparative statics under identical tax rates $t_i = t$ ($i = 1,2$):

$$\frac{\partial g_{i3}}{\partial t_i} = \frac{\partial g_{j3}}{\partial t_i} = \frac{\mu}{2\delta} > 0, \qquad i, j = 1,2, \tag{B-5}$$

$$\frac{\partial y_i}{\partial t_i} = -\frac{y[1-\mu c'(y)]+[y-\mu c(y)-\mu^2 t/\delta]}{2y(1-\mu t)c''(y)} < 0, \qquad i = 1,2, \tag{B-6}$$

$$\frac{\partial y_j}{\partial t_i} = \frac{\mu[yc'(y)-c(y)-\mu t/\delta]}{2y(1-\mu t)c''(y)}, \quad i,j = 1,2;\ i \neq j. \tag{B-7}$$

Using (26), (B-6) and (B-7), it is straightforward to show: $\forall \mu \in [0,1)$, $\frac{\partial y_i}{\partial t_i} - \frac{\partial y_j}{\partial t_i} = -\frac{y-\mu c(y)-\mu^2 t/\delta}{y(1-\mu t)c''(y)} < 0$ and $\frac{\partial y_i}{\partial t_i} + \frac{\partial y_j}{\partial t_i} = -\frac{1-\mu}{(1-\mu t)^2 c''(y)} < 0$, which implies $\left|\frac{\partial y_i}{\partial t_i}\right| > \left|\frac{\partial y_j}{\partial t_i}\right|$. Hence, when costs are not fully deductible, the own effect of country i's tax rate on affiliate i's output is stronger than the cross-border effect on affiliate j's output. Furthermore, we can derive that: $\forall \mu \in [0,1]$, $\frac{\partial \lambda_i}{\partial t_i} = -\frac{\partial \lambda_j}{\partial t_i} = -\frac{y-\mu c(y)-\mu^2 t/\delta}{4y^2(1-\mu t)c''(y)} < 0$. As discussed in Subsection 3.2, the MNE manipulates the sales formula by favorably adjusting output levels in FA countries. It raises (reduces) output in the FA country with a lower (higher) tax rate, through which it can increase the share of the consolidated tax base assigned to the low-tax country and reduce its tax liability.

In the first stage, country i chooses tax rate t_i to maximize its revenue $R_i(t_i,t_j) = t_i \lambda_i \pi^c$, where π^c is given by (B-1). In doing so, it takes into account the MNE's behavior that depends on countries' tax rates due to (B-2) – (B-4). Solving the first-order conditions of revenue maximization problem and using the symmetry property yields the following:

$$\frac{\partial R_i}{\partial t_i} = y - \mu c(y) - \frac{3\mu^2 t}{2\delta} - \frac{t\left[y-\mu c(y)-\frac{\mu^2 t}{\delta}\right]^2}{2y^2(1-\mu t)c''(y)} - \frac{t(1-\mu)^2}{2(1-\mu t)^3 c''(y)} = 0, \tag{B-8}$$

where y is determined by (26) (with country subscripts dropped due to symmetry).

Given the specific functional form of $c(y)$, solving (B-8) yields each FA country's Nash equilibrium tax rate $\hat{t}(\delta,\mu)$. Similar to SA, (26) indicates that the equilibrium output $\hat{y}$ in each FA country directly depends on the equilibrium tax $\hat{t}$ and deductibility parameter μ, i.e., $\hat{y} = \hat{y}(\hat{t},\mu)$. Then given (10), (B-1) and (B-3), each FA country's equilibrium revenue is still characterized by (31), except that tildes are replaced by hats.

The following proposition is the FA analog to Proposition 3(i), showing the desirability of reducing tax deductibility under low tax capacity.

Proposition B.1. *Under FA, a tax regime with lower deductibility can yield a larger tax revenue when tax capacity is sufficiently low. That is,* $\forall \mu_1, \mu_2$ *with* $0 \le \mu_1 < \mu_2 \le 1$, *there exists some* $\delta' > 0$ *such that* $\hat{R}(\delta, \mu_1) > \hat{R}(\delta, \mu_2)$ *holds when* $\delta < \delta'$.

Proof. The reasoning in the proof of Proposition 3(i) also applies to FA. Using (26), (28) and (31) (with the tildes replaced by hats), we can compute the effect of tax capacity on the equilibrium tax revenue under FA:

$$\frac{\partial \hat{R}(\delta,\mu)}{\partial \delta} = \frac{\partial(\hat{t}\hat{\pi}^t(\hat{t},\delta,\mu))}{\partial \hat{t}} \cdot \frac{\partial \hat{t}(\delta,\mu)}{\partial \delta} + \hat{t} \cdot \frac{\partial \hat{\pi}^t(\hat{t},\delta,\mu)}{\partial \delta}, \tag{B-9}$$

where

$$\frac{\partial(\hat{t}\hat{\pi}^t(\hat{t},\delta,\mu))}{\partial \hat{t}} = \hat{y} - \mu c(\hat{y}) - \frac{2\mu^2 \hat{t}}{\delta} - \frac{\hat{t}(1-\mu)^2}{(1-\mu\hat{t})^3 c''(\hat{y})} \quad \text{and} \quad \frac{\partial \hat{\pi}^t(\hat{t},\delta,\mu)}{\partial \delta} = \frac{\mu^2 \hat{t}}{\delta^2} > 0.$$

Given (26) and (B-8), we can obtain the following:

$$\lim_{\delta \to 0} \hat{t}(\delta,\mu) = 0, \quad \lim_{\delta \to 0} \hat{y}(\delta,\mu) = \bar{y}, \quad \text{and} \quad \lim_{\delta \to 0} \frac{\hat{t}(\delta,\mu)}{\delta} = \lim_{\delta \to 0} \frac{\partial \hat{t}(\delta,\mu)}{\partial \delta} = \frac{2[\bar{y} - \mu c(\bar{y})]}{3\mu^2},$$

$$\forall \mu \in (0,1], \tag{B-10}$$

where the first equality in the third part of (B-10) holds by L'Hôpital's rule.

Taking the limit on both sides of (B-9) as $\delta \to 0$ and using (B-10) yields:

$$\lim_{\delta \to 0} \frac{\partial(\hat{t}\hat{\pi}^t(\hat{t},\delta,\mu))}{\partial \hat{t}} = -\frac{\bar{y} - \mu c(\bar{y})}{3} < 0, \quad \lim_{\delta \to 0}\left(\hat{t} \cdot \frac{\partial \hat{\pi}^t(\hat{t},\delta,\mu)}{\partial \delta}\right) = \frac{4[\bar{y} - \mu c(\bar{y})]^2}{9\mu^2}, \quad \text{and}$$

$$\lim_{\delta \to 0} \frac{\partial \hat{R}(\delta,\mu)}{\partial \delta} = \frac{2[\bar{y} - \mu c(\bar{y})]^2}{9\mu^2}, \quad \forall \mu \in (0,1]. \tag{B-11}$$

(A-28) and the third part of (B-11) indicate that as $\delta \to 0$, the effect of tax capacity on the equilibrium revenue under SA and FA converges in the same way. Thus, we have:

$\forall \mu_1, \mu_2 \in [0,1]$ with $\mu_1 < \mu_2$, there exists some $\delta' > 0$ such that $\hat{R}(\delta, \mu_1) > \hat{R}(\delta, \mu_2)$ when $0 < \delta < \delta'$. **Q.E.D.**

Our analysis reveals that the tax competition can yield very different fiscal externalities under SA and FA. Specifically, (A-24) indicates that the fiscal externality under SA is positive and causes inefficient undertaxation. In sharp contrast, the first part of (B-11) implies that with sufficiently low tax capacity the fiscal externality under FA is negative and hence results in inefficient overtaxation. This relates our paper to Riedel and Runkel (2007). In the case where countries maximize tax revenues and costs are fully tax deductible, they argue that there is overtaxation under FA if the water's edge externality dominates the formula externality. Our paper shows that for any given deductibility parameter $\mu \in (0,1]$, overtaxation will occur in the FA union when countries' tax capacity is sufficiently low.

Appendix C. Comparing profit and turnover taxation in general environments

In this appendix, we drop the specific form of the production cost function (i.e., a linear-power function) in the main text. Instead, we assume a general form of $c(y)$ that has the following properties:

(i) $c(0)=0$; (ii) $0 \le c'(0) < 1$, $c'(+\infty) > 1$; (iii) $c''(y) > 0$, $2c''(y) + yc'''(y) > 0$, $\forall y > 0$; (iv) $sign\{c'''(y)\} = \text{constant}$, $\forall y > 0$. (C-1)

Assumption (i) in (C-1) means zero production cost for zero output. Combining (ii) and the first (second) part of (iii) guarantees a unique equilibrium output level under profit taxation (turnover taxation). Assumption (iv) indicates that the marginal production cost function (i.e., the inverse supply curve of the MNE) can be either concave or convex, whereas the concavity or convexity is non-reversal. In the case of $c'''(.) \le 0$, the turnover tax, as shown by (C-7) below, can cause a large loss of production efficiency. In the case of $c'''(.) > 0$, the value of $c''(\overline{y})$ is high so that the formula manipulation effect under FA is mild. So in both cases, the adverse effect of formula manipulation is minor compared to the gain in production efficiency when countries switch from turnover tax to FA. It is then intuitive that FA is superior to turnover tax under a sufficiently high tax capacity. Notably, the production cost production specified in the main text satisfies all the four assumptions in (C-1).

In what follows, we formally prove that the comparison of profit and turnover taxation characterized in Proposition 1(i) holds for the general form of the production technology. Firstly, the lemma below provides the equilibrium outcomes under turnover taxation.

Lemma C.1. *The equilibrium tax rate and revenue of each country under turnover tax are:*
$t^T = \underline{y}c''(\underline{y})$ *and* $R^T = \underline{y}[1-c'(\underline{y})] = \underline{y}^2 c''(\underline{y})$, (C-2)
where equilibrium output $\underline{y}$ *is implicitly determined by* $c'(\underline{y}) + \underline{y}c''(\underline{y}) = 1$.

Proof. Using (20), we can rewrite country i's tax revenue under turnover tax as a function of its output, i.e., $R_i = y_i[1-c'(y_i)]$. Then solving the first-order condition of the revenue maximization yields:
$c'(\underline{y}) + \underline{y}c''(\underline{y}) = 1$. (C-3)

Given (20) and (C-3), (C-2) immediately follows. **Q.E.D.**

Recall from (5) and (13) that the efficient output level of each country under profit tax is characterized by $c'(\overline{y}) = 1$. In (C-3), $\underline{y}c''(\underline{y}) > 0$ represents the "production wedge", which makes the output under turnover tax smaller than under profit tax, i.e., $\underline{y} < \overline{y}$.

Notice that under profit tax with SA the Nash equilibrium tax rate and revenue of the two countries are still characterized by (7) and (8), respectively. Under profit tax with FA, the Nash equilibrium tax of the FA countries is still determined by (17). In terms of each country's equilibrium tax revenue, we can compare profit and turnover taxation in the limit cases:

Lemma C.2. *Turnover tax is superior (inferior) to profit tax as tax capacity approaches zero (infinity): (i)* $\lim_{\delta\to 0}\tilde{R}(\delta) < R^T$ *and* $\lim_{\delta\to 0}\hat{R}(\delta) < R^T$ *; (ii)* $\lim_{\delta\to +\infty}\tilde{R}(\delta) > R^T$ *and* $\lim_{\delta\to +\infty}\hat{R}(\delta) > R^T$.

Proof. It is obvious from (8) and (C-2) that: $\lim_{\delta\to 0} R^{SA}(\delta) = 0 < R^T$. As for FA, it follows from (A-2) that $\lim_{\delta\to 0}\hat{g} = \frac{2\bar{\pi}}{3}$. Then we can get: $\lim_{\delta\to 0}\hat{R}(\delta) = \lim_{\delta\to 0}(\delta\hat{g})(\bar{\pi}-\hat{g}) = 0 < R^T$.

We now prove (ii). Given (8) and (C-2), it is straightforward to show:

$$\lim_{\delta\to +\infty}\tilde{R}(\delta) = \bar{\pi} = \int_0^{\bar{y}}[1-c'(y)]dy > \int_0^{\underline{y}}[1-c'(y)]dy > \int_0^{\underline{y}}[1-c'(y)-yc''(y)]dy = R^T .$$

On the other hand, taking the limit on both sides of (17) as $\delta\to+\infty$, we can derive:

$$\lim_{\delta\to +\infty}\left[\bar{\pi}-\frac{3\hat{t}}{2\delta}-\frac{\hat{t}(\bar{\pi}-\frac{\hat{t}}{\delta})^2}{2(1-\hat{t})\bar{y}^2c''(\bar{y})}\right] = \bar{\pi}-\frac{\bar{\pi}^2(\lim_{\delta\to+\infty}\hat{t})}{2(1-\lim_{\delta\to+\infty}\hat{t})\bar{y}^2c''(\bar{y})} = 0 \;\Rightarrow$$

$$\lim_{\delta\to +\infty}\hat{t} = \frac{2\bar{y}^2c''(\bar{y})}{\bar{\pi}+2\bar{y}^2c''(\bar{y})} \;\Rightarrow\; \lim_{\delta\to +\infty}\hat{R}(\delta) = \lim_{\delta\to +\infty}\hat{t}(\bar{\pi}-\frac{\hat{t}}{\delta}) = \frac{2\bar{\pi}\bar{y}^2c''(\bar{y})}{\bar{\pi}+2\bar{y}^2c''(\bar{y})}. \tag{C-4}$$

In the following, we investigate two cases: (i) $c'''(.)\le 0$ and (ii) $c'''(.)>0$, separately.

Case (i): $c'''(y)\le 0$, $\forall y>0$.

From the second part of assumption (iii) in (C-1), $y^2c''(y)$ is strictly increasing in y. Then it follows from $\bar{y}>\underline{y}$ that:

$$\bar{y}^2c''(\bar{y}) > \underline{y}^2c''(\underline{y}) = R^T . \tag{C-5}$$

Besides, notice that:

$$\forall y>0,\; 1-c'(y) \ge -c''(\underline{y})(y-\underline{y})+1-c'(\underline{y}) = (2\underline{y}-y)c''(\underline{y}), \tag{C-6}$$

where the first inequality derives from the concavity of function $c'(.)$, and the last equality holds due to (C-3).

Letting $y=\bar{y}$ in (C-6), we get: $1-c'(\bar{y}) = 0 \ge (2\underline{y}-\bar{y})c''(\underline{y})$, which leads to:

$$\bar{y}\ge 2\underline{y} . \tag{C-7}$$

Given (C-2), (C-6) and (C-7), it is straightforward to show:

$$\bar{\pi} = \int_0^{\bar{y}}[1-c'(y)]dy \ge \int_0^{2\underline{y}}[1-c'(y)]dy \ge \int_0^{2\underline{y}}(2\underline{y}-y)c''(\underline{y})dy = 2\underline{y}^2c''(\underline{y}) = 2R^T . \tag{C-8}$$

Combining (C-4), (C-5) and (C-8) yields:

$$\lim_{\delta\to +\infty}\hat{R}(\delta) = \frac{2\bar{\pi}\bar{y}^2c''(\bar{y})}{\bar{\pi}+2\bar{y}^2c''(\bar{y})} > \frac{2\cdot 2R^T\cdot R^T}{2R^T+2R^T} = R^T .$$

Case (ii): $c'''(y)>0$, $\forall y>0$.

Firstly, notice that:

$$\forall\theta\in(0,1),\ 1-c'((1-\theta)\underline{y}+\theta\overline{y})>1-(1-\theta)c'(\underline{y})-\theta c'(\overline{y})=(1-\theta)[1-c'(\underline{y})], \qquad \text{(C-9)}$$

where the first inequality holds since $c'(.)$ is strictly convex, and the last equality holds due to $c'(\overline{y})=1$.

Then we can show the following:

$$\begin{aligned}\overline{\pi} &= \int_0^{\underline{y}}[1-c'(y)]dy+\int_{\underline{y}}^{\overline{y}}[1-c'(y)]dy \\ &= \int_0^{\underline{y}}[1-c'(y)]dy+\int_0^1(\overline{y}-\underline{y})[1-c'((1-\theta)\underline{y}+\theta\overline{y})]d\theta \\ &> \int_0^{\underline{y}}[1-c'(\underline{y})]dy+\int_0^1(\overline{y}-\underline{y})(1-\theta)[1-c'(\underline{y})]d\theta \\ &= \frac{\underline{y}+\overline{y}}{2}[1-c'(\underline{y})]=\frac{1}{2}(\underline{y}+\overline{y})\underline{y}c''(\underline{y})\end{aligned}, \qquad \text{(C-10)}$$

where the first inequality follows from $c''(.)>0$ and (C-9), and the last equality is obtained by using (C-3).

The strict convexity of function $c'(.)$ implies $c'(\underline{y})>c''(\overline{y})(\underline{y}-\overline{y})+c'(\overline{y})$, from which we obtain:

$$\overline{y}^2c''(\overline{y})>\overline{y}^2\frac{c'(\overline{y})-c'(\underline{y})}{\overline{y}-\underline{y}}=\frac{\overline{y}^2\underline{y}c''(\underline{y})}{\overline{y}-\underline{y}}, \qquad \text{(C-11)}$$

where the last equality is obtained by $c'(\overline{y})=1$ and (C-3).

Combining (C-4), (C-10) and (C-11) leads to:

$$\lim_{\delta\to+\infty}\hat{R}(\delta)=\frac{2\overline{\pi}\overline{y}^2c''(\overline{y})}{\overline{\pi}+2\overline{y}^2c''(\overline{y})}>\frac{2[(\underline{y}+\overline{y})\underline{y}c''(\underline{y})/2]\cdot[\overline{y}^2\underline{y}c''(\underline{y})/(\overline{y}-\underline{y})]}{(\underline{y}+\overline{y})\underline{y}c''(\underline{y})/2+2\overline{y}^2\underline{y}c''(\underline{y})/(\overline{y}-\underline{y})}=\frac{2\overline{y}^2\underline{y}(\underline{y}+\overline{y})c''(\underline{y})}{5\overline{y}^2-\underline{y}^2},$$

from which we have:

$$\lim_{\delta\to+\infty}\hat{R}(\delta)-R^T>\frac{2\overline{y}^2\underline{y}(\underline{y}+\overline{y})c''(\underline{y})}{5\overline{y}^2-\underline{y}^2}-\underline{y}^2c''(\underline{y})=\frac{\underline{y}(\overline{y}-\underline{y})^2(2\overline{y}+\underline{y})c''(\underline{y})}{5\overline{y}^2-\underline{y}^2}>0.$$

Therefore, in both cases, $\lim_{\delta->+\infty}\hat{R}(\delta)>R^T$ holds. **Q.E.D.**

According to Lemma 1(iii) and Lemma 2(iii), both $\tilde{R}(\delta)$ and $\hat{R}(\delta)$ are continuous and strictly increasing in δ. Then together with Lemma C.2, it is obvious that Proposition 1(i) holds with the general form of production technology.

References

ActionAid, 2012. *Calling Time: Why SABMiller Should Stop Dodging Taxes in Africa*. London: ActionAid.

ActionAid, 2015. *An Extractive Affair: How one Australian Mining Company's Tax Dealings are Costing the World's Poorest Country Millions*. London: ActionAid.

Agrawal, D. R., Wildasin, D. E., 2020. Technology and tax systems. *Journal of Public Economics* 185, 104082.

Albertin, G., Yontcheva, B., Devlin, D., Devine, H., Gerard, M., Suljagic, I., Thakoor, V., Beer, S., 2021. Tax Avoidance in Sub-Saharan Africa's Mining Sector. *IMF Departmental Paper* 2021/022.

Auerbach, A. J., Devereux, M. P., 2018. Cash-flow taxes in an international setting. *American Economic Journal: Economic Policy* 10(3), 69-94.

Best, M. C., Brockmeyer, A., Kleven, H. J., Spinnewijn, J., Waseem, M., 2015. Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan. *Journal of Political Economy* 123(6), 1311-1355.

Bilicka, K. A., 2019. Comparing UK tax returns of foreign multinationals to matched domestic firms. *American Economic Review* 109(8), 2921-2953.

Bunn, D., Asen, E., Enache, C., 2020. *Digital Taxation around the World.* Washington, DC: Tax Foundation.

Clausing, K. A., 2016. The effect of profit shifting on the corporate tax base in the United States and beyond. *National Tax Journal* 69(4), 905–934.

Crivelli, E., De Mooij, R., Keen, M., 2016. Base erosion, profit shifting and developing countries. *FinanzArchiv: Public Finance Analysis* 72(3), 268-301.

Davies, R. B., Martin, J., Parenti, M., Toubal, F., 2018. Knocking on tax haven's door: multinational firms and transfer pricing. *Review of Economics and Statistics* 100(1), 120-134.

Devereux, M. P., Lockwood, B., Redoano, M., 2008. Do countries compete over corporate tax rates? *Journal of Public Economics* 92(5-6), 1210-1235.

Diamond, P. A., Mirrlees, J. A., 1971. Optimal taxation and public production I: production efficiency. *American Economic Review* 61(1), 8-27.

Dischinger, M., Riedel, N., 2011. Corporate taxes and the location of intangible assets within multinational firms. *Journal of Public Economics* 95(7-8), 691-707.

Eichner, T., Runkel, M., 2008. Why the European Union should adopt formula apportionment with a sales factor. *Scandinavian Journal of Economics* 110(3), 567-589.

Eichner, T., Runkel, M., 2011. Corporate income taxation of multinationals in a general equilibrium model. *Journal of Public Economics* 95(7-8), 723-733.

Ernst & Young, 2015. Experiences with cash-flow taxation and prospects. *European Commission Taxation Papers, Working Paper* 55–2015.

European Commission, 2016. *Proposal for a Council Directive on a Common Consolidated Corporate Tax Base (CCCTB)*, COM (2016) 683 final.

European Commission, 2017. *Commission decision of 4.10.2017 on state aid SA.38944 (2014/EC) (ex 2014/NN) implemented by Luxembourg to Amazon*, C(2017) 6740 final.

European Commission, 2018. *Proposal for a Council Directive on the Common System of a Digital Services Tax on Revenues Resulting from the Provision of Certain Digital Services*, COM (2018) 148 final.

Gaspar Committee, 2014. *Report of the Commission Expert Group on Taxation of the Digital Economy*, chaired by Vitor Gaspar, Brussels: European Commission.

Gumpert, A., Hines Jr, J. R., Schnitzer, M., 2016. Multinational firms and tax havens. *Review of Economics and Statistics* 98(4), 713-727.

Gresik, T. A., 2016. Allowing firms to choose between separate accounting and formula apportionment taxation. *Journal of Public Economics* 138, 32-42.

Hines, J. R., 2005. Do Tax Havens Flourish? In J. M. Poterba, ed. *Tax Policy and the Economy*, vol. 19, 65–99. Cambridge, MA: MIT Press.

Johannesen, N., 2010. Imperfect tax competition for profits, asymmetric equilibrium and beneficial tax havens. *Journal of International Economics* 81(2), 253-264.

Keen, M., Mullins, P., 2017. International corporate taxation and the extractive industries: principles, practice, problems. In P. Daniel, M. Keen, A. Swistak and V. Thuronyi (eds), *International Taxation and the Extractive Industries*. New York and London: Routledge.

Keen, M., Wildasin, D., 2004. Pareto-efficient international taxation. *American Economic Review* 94(1), 259-275.

Kind, H. J., Midelfart, K. H., Schjelderup, G., 2005. Corporate tax systems, multinational enterprises, and economic integration. *Journal of International Economics* 65(2), 507-521.

Klein, D., Ludwig, C. A., Spengel, C., 2022. Taxing the digital economy: Investor reaction to the european commission's digital tax proposals. *National Tax Journal 75*(1), 61-92.

Krautheim, S., Schmidt-Eisenlohr, T., 2011. Heterogeneous firms, 'profit shifting' FDI and international tax competition. *Journal of Public Economics* 95(1-2), 122-133.

Mardan, M., Stimmelmayr, M., 2018. Tax revenue losses through cross-border loss offset: an insurmountable hurdle for formula apportionment? *European Economic Review* 102, 188-210.

Munk, K. J., 1980. Optimal taxation with some non-taxable commodities. *The Review of Economic Studies* 47(4), 755-765.

Nielsen, S. B., Raimondos-Møller, P., Schjelderup, G., 2010. Company taxation and tax spillovers: separate accounting versus formula apportionment. *European Economic Review* 54(1), 121-132.

OECD, 2018. *Tax Challenges Arising from Digitalisation – Interim Report 2018*, Paris: OECD Publishing.

Palmer, J. A., 2003. Relative convexity. *ECE Dept., UCSD, Tech. Rep.*

Riedel, N., Runkel, M., 2007. Company tax reform with a water's edge. *Journal of Public Economics* 91(7-8), 1533-1554.

Slemrod, J., Wilson, J. D., 2009. Tax competition with parasitic tax havens. *Journal of Public Economics* 93(11-12), 1261-1270.

Stiglitz, J. E., Dasgupta, P., 1971. Differential taxation, public goods, and economic efficiency. *The Review of Economic Studies* 38(2), 151-174.

Tørsløv, T., Wier, L., Zucman, G., 2022. The missing profits of nations. *The Review of Economic Studies*. Forthcoming.

Vives, X., 2000. *Oligopoly Pricing: Old Ideas and New Tools.* MIT Press, Cambridge, Massachusetts.

Zucman, G., 2014. Taxing across borders: tracking personal wealth and corporate profits. *Journal of Economic Perspectives 28*(4), 121-148.