

Three Essays on International Taxation of Multinational Enterprises

Doctoral Dissertation Presented by

Xuyang CHEN

In Fulfillment of the Requirements for Joint Doctoral Degree in Economics and Management
at Université catholique de Louvain & Université libre de Bruxelles

Thesis Committee:

Jean Hindriks (Supervisor), UCLouvain

Mathieu Parenti (Co-supervisor), ULB & Paris School of Economics

Luca Merlino (Chairman), ULB

Gonzague Vanoorrenberghe (Secretary), UCLouvain

Hylke Vandenbussche, KU Leuven

Andreas Haufler, University of Munich



October, 2024

Acknowledgments

This doctoral thesis is the result of several years of study and research, and it would not have been possible without the support and guidance of many people, to whom I am deeply grateful.

First, I would like to express my gratitude to my supervisor, Prof. Jean Hindriks, for introducing me to the very important and interesting research area: taxation of multinational firms. He provided guidance throughout my doctoral studies, and I benefited a lot from our discussions and his insightful comments on my research. I am indebted to him for his encouragement, patience, and support. I am also deeply thankful to my co-supervisor, Prof. Mathieu Parenti, for his valuable comments and suggestions on my work (particularly on Chapter 3).

I would like to thank the members of my jury: Prof. Andreas Haufler, Prof. Luca Merlino, Prof. Gonzague Vanoorrenberghe and Prof. Hylke Vandenbussche. I greatly appreciate their constructive comments and feedbacks, which helped me to improve the quality of this thesis.

I also sincerely thank Prof. Paul Belleflamme for his kindness and for helpful discussions and comments at the PROSEco meeting. I joined his research group when working on Chapter 2, which relates to some tax policies for the digital economy. Besides, I would like to express my gratitude to Ms. Catherine Germain and Ms. Anne-Marie Notarianni, who were always available to assist me with various administrative matters.

I gratefully acknowledge the financial support from ARC grant (No. 19/24-101), co-financement COVID, ARC grant (4110S000001), and FNRS grant (MIS 4129S000012). The funding was essential in supporting my research and making this thesis possible.

Finally, I wish to thank my family for their unconditional love and support throughout the doctoral journey. In particular, I am deeply grateful to my parents for always supporting my decision, even when it meant that I would be living abroad for several years.

Contents

Introduction.....	1
1. Tax Enforcement Leadership and Cooperation Against Profit Shifting	7
1.1. Introduction	8
1.2. The model.....	10
1.2.1. The setups	11
1.2.2. Enforcement leadership	12
1.2.3. The equilibrium analysis.....	13
1.3. Cooperative enforcement	19
1.3.1. Nash bargaining over enforcement.....	19
1.3.2. Bargaining solution	20
1.4. Considering tax leadership	23
1.5. Conclusion.....	24
Appendix 1.A. Proofs.	25
2. Multinational Taxation Under Pressure: The Role of Tax Deductibility.....	35
2.1. Introduction	36
2.2. The model.....	39
2.3. Corporate tax regimes	41
2.3.1. Profit tax with separate accounting	41
2.3.2. Profit tax with formula apportionment	43
2.3.3. Turnover taxation	46
2.4. Comparison of the tax regimes	49
2.4.1. Profit versus turnover taxation	49
2.4.2. Comparing the three tax regimes.....	50
2.5. The general case of tax deductibility.....	53
2.6. Conclusion.....	59
Appendix 2.A. Proofs.	60
Appendix 2.B. The impact of tax deductibility under FA.....	70
Appendix 2.C. Comparing profit and turnover taxation in general environments	74
3. The Global Minimum Tax, Investment Incentives and Asymmetric Tax Competition	78

3.1. Introduction	79
3.2. The model.....	84
3.3. Short-run analysis of the GMT: fixed tax rates	87
3.4. Long-run analysis: tax competition.....	92
3.4.1. The equilibrium analysis.....	92
3.4.2. The revenue effects of the GMT	99
3.5. Conclusion.....	102
Appendix 3.A. Proofs and derivations	103
Appendix 3.B. Conditions for excess profits to be positive	123
Appendix 3.C. Sufficient conditions for the quasiconcavity of $R_2^m(t_m)$, $\forall t_m \in (t_2^N, t_1^N)$	124
Appendix 3.D. The analysis of the case of $\sigma \in (0, \underline{\sigma}]$	126
Appendix 3.E. The extension: including labor	130
General Conclusion: Lessons for International Taxation of MNEs.....	134
References.....	136

Introduction

Multinational enterprises (MNEs) are prospering. Multinational profits grew by more than 50%, from \$1703 billion in 2015 to \$2509 billion in 2019. The growth in multinational profits also outpaced the growth in global corporate profits, so that the share of multinational profits in corporate profits increased from 15% to 18% over the period 2015 – 2019 (see Wier and Zucman, 2022). MNEs have brought managerial practices, innovation, investment, increased productivity and plenty of job opportunities to host countries, which largely contribute to local economic development. Nevertheless, MNEs also have more avenues to avoid paying taxes, compared to domestic firms that do not have access to international networks. A number of studies show that MNEs are reducing tax liabilities by shifting profits from high-tax countries to low-tax countries and that profit shifting has caused considerable losses of tax revenues for both OECD and developing countries. International profit shifting by MNEs has been at the forefront of policy debates and news coverage in recent years.

In practice, MNEs shift profits between non-tax-haven countries and to tax havens, while the latter case causes significantly larger losses of tax revenues. It is estimated that \$969 billion in profits were shifted to tax havens globally in 2019, the equivalent of 37.4% of global multinational profits. Accordingly, foreign firms in tax havens usually display much higher profitability (captured by the ratio of pre-tax profits to wages) than local firms. For example, in 2019, Puerto Rico had an exceptionally high profits-to-wage ratio of 1686% for foreign firms. In Luxembourg and Ireland, the profits-to-wage ratios of foreign firms were 600% and 487%, respectively. In absolute terms, Singapore was the world's largest recipient of shifted profits in 2019. It attracted \$132.3 billion of foreign profits, compared to \$129.6 billion for Ireland (see Wier and Zucman, 2022). One prominent method used for profit shifting is manipulating transfer prices of intra-firm transactions. Specifically, high-tax affiliates can sell (buy) goods or services to (from) lower-tax affiliates at artificially low (high) prices, which lowers the total tax liability of MNEs. Another method used by MNEs to shift profits involves internal loans. For most countries' tax systems, debt and equity are treated differently for tax purposes, as interest payments are tax deductible. This creates an incentive for financing with debt rather than equity. In practice, low-tax affiliates can lend a loan to high-tax affiliates, which enables the latter to reduce tax payments by deducting interest payments from their taxable incomes. As a key consequence of profit shifting, countries compete with each other in corporate tax rates to attract multinational profits, and the tax competition has resulted in substantial falls in corporate tax rates. In 1980, the average worldwide statutory tax rate was 40.11%. In 2020, the average statutory rate was 23.85%, representing a 40.5% decrease over the forty years. The eroded tax bases together with reduced tax rates lead to enormous revenue losses worldwide.

In a globalized and digitalized world, how to effectively tax MNEs and secure revenues has been of first-order importance for many governments. **This thesis focuses on several developments of anti-avoidance rules and policies, and theoretically investigates their effects on MNEs' behavior and countries' corporate taxes in the models of international tax competition.** The thesis is structured into three chapters, discussing how to tackle profit shifting and improve taxation of multinationals from different perspectives. A common feature of the three chapters is that the costs associated with MNEs' profit shifting activities are assumed to be increasing in countries' tax capacity (enforcement). Indeed, MNEs have to spend more legal resources to defend their transfer pricing process and may face higher expected penalties under higher tax capacities. On the other hand, Chapter 1 endogenizes countries' tax capacity (enforcement) while assuming a profit tax with full deductibility. Chapters 2 and 3 assume the exogeneity of tax capacity, but allow for partial deductibility to consider the effects of real economic activities of MNEs. Such treatment makes the models tractable and enables us to concentrate on the specific tax issue analyzed in each chapter. In what follows, I briefly introduce the three chapters, respectively.

Chapter 1 is related to building tax capacity and coordinating tax enforcements. Countries worldwide are spending more resources enforcing the rules that govern the taxation of MNEs. Empirical studies document a trend towards stricter anti-tax avoidance rules and policies over time. The measures that countries take to enhance tax enforcement include tightening transfer pricing regulations, increasing the auditing probability and designing effective controlled-foreign-company rules. In particular, the OECD and G20 in 2015 launched a base erosion and profit shifting (BEPS) project which provides 15 action plans against tax avoidance. Over 145 countries and jurisdictions now actively participate in the development of this OECD/G20 BEPS project. On the other hand, evidence suggests that countries' decisions on enhancing tax enforcement are not made simultaneously (see Zinn et al., 2014; Riedel et al., 2015). In response to the OECD/G20 BEPS project, the progresses of enforcement reforms are also different among the participating countries.

In light of these facts, this chapter considers a fiscal competition game – where profit shifting levels, tax rates and enforcements are endogenously determined – to study tax enforcement leadership and cooperation. To focus on tax enforcement issues and keep the model tractable, we assume that countries levy a profit tax with costs fully deductible. In this setup, the productions and prices are undistorted, implying that consumer surplus is independent of the tax system. Accordingly, in this chapter we refer to the tax revenue net of the enforcement cost as a country's welfare. We will relax the assumption of full tax deductibility to consider the effects of real economic activities in Chapters 2 and 3, while the tax capacity (enforcement) is assumed to be exogenous in the last two chapters.

The key novelty of our setting is twofold. First, our model incorporates two types of asymmetry between countries. Besides the asymmetry in market size, countries are also asymmetric in enforcement productivity (captured by the enforcement elasticity of tax revenue). Second, the timing of enforcement decisions is endogenized. To this end, we

introduce a timing game at the pre-play stage, in which countries commit themselves to move early or late in choosing tax enforcements. Given the timing decisions at the pre-play stage, countries set their enforcement levels to maximize welfare before competing in corporate tax rates. The MNEs chooses output and the reported profit in each country to maximize total after-tax profits, taking as given the tax environments. Then we further study the international enforcement cooperation, which is described as a Nash bargaining problem. Specifically, countries negotiate over enforcement levels to maximize the product of their welfare gains relative to equilibrium outcomes of the enforcement leadership game.

The main results of this chapter are as follows. The subgame perfect equilibria of the game are two Stackelberg equilibria with countries setting tax enforcements sequentially. Compared to the simultaneous Nash equilibrium, the Stackelberg equilibria can induce higher enforcements and welfare for both countries. By the risk-dominance criterion, the low-enforcement productivity country would be the leader. So we present a novel driving force of enforcement leadership (the asymmetric enforcement productivity), which is different from that of tax leadership (the asymmetric market size). However, tax enforcements are still inefficiently low at Stackelberg equilibria, and the positive externality of enforcement cannot be fully internalized. We establish that Nash bargaining over enforcements (with countries competing in taxes) can induce higher enforcement, tax and revenue for each country. Furthermore, countries' cooperative welfare does not increase in the same proportion. Relative to the Stackelberg outcome, the low-enforcement productivity country can benefit more from enforcement cooperation. Our results are robust to the endogenous tax leadership.

Chapter 2 is related to limiting tax deductibility, i.e., the share of costs that can be deducted from a firm's gross revenue. Tax policy developments show that a number of countries are limiting the deductibility of costs. In digital sectors, some countries are using turnover taxes targeted at gross revenues from digital products and services. Examples include India's Equalisation Levy, Italy's levy on digital transactions, Hungary's advertisement tax and France's tax on revenues from provision of digital interfaces and advertising services (see OECD, 2018). Moreover, the European Commission proposed an interim digital services taxes (DSTs) with a rate of 3% on gross revenues in the European Union from specific digital services characterized by user value creation (see European Commission, 2018). 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. Compared to income-based taxes, royalties are usually levied on the gross value of the product to reduce vulnerabilities to profit shifting, and hence they are akin to turnover taxes. Besides, nine of the 15 resource-intensive African countries implement alternative minimum taxes (AMTs) in the mining sector to ensure some corporate taxes are paid each year (see Albertin et al., 2021). Under AMTs, firms are subject to the corporate income tax or turnover tax, 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. Nevertheless, deviating from profit taxation by limiting the tax deductibility results in production inefficiency. Critics argue that turnover taxes should, by and large, be removed to avoid the distortions that taxes on revenues create.

In this chapter, we take into account MNEs' production technology and countries' tax capacity to investigate the impact of tax deductibility on MNEs' behavior and countries' corporate taxes. To this end, we set up a formal tax competition model, in which two countries set corporate tax rates 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 are either fully tax deductible (i.e., a pure profit tax) or fully non-deductible (i.e., 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). Different from SA under which profit is taxed in the country where the MNE declares it, under FA two countries' tax bases are consolidated first and then assigned to each country through a sales-based formula. We identify the optimal choice among the three tax regimes from the perspective of equilibrium tax revenues. Then we analyze the general case of tax deductibility where countries' tax bases can range from a pure profit tax base to a turnover base, and revisit the Diamond-Mirrlees production efficiency theorem (Diamond and Mirrlees, 1971) in international profit shifting context.

The main results of Chapter 2 are as follows. A profit tax with either SA or FA can preserve efficient production at the expense of tax base erosion. In particular, although a profit tax with FA eliminates 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 production distortion. 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. The tax revenue comparisons of the three tax regimes depend on the MNE's production technology and countries' tax capacity. Specifically, profit tax with FA dominates when the output price elasticity is high and the tax capacity is intermediate. Otherwise, turnover tax dominates for low tax capacity, and profit tax with SA dominates for high tax capacity. In the general case of tax deductibility, we show that the validity of the Diamond-Mirrlees theorem in international profit shifting context crucially depends on countries' tax capacity and industry-specific production technologies. Specifically, a lower deductibility can increase tax revenue under sufficiently low tax capacity. By contrast, equilibrium taxes strictly increase with tax deductibility when tax capacity is high, marginal production cost is low or the output elasticity is high. Numerical simulations further indicate that the optimal tax deductibility is increasing in tax capacity and output elasticity, and decreasing in marginal production cost.

Chapter 3 is related to the OECD's global minimum tax (GMT). In 2021, over 135 jurisdictions agreed on a 15% GMT, which is a key part of Pillar Two of the two-pillar

solution proposed by the OECD/G20 Inclusive Framework on BEPS. The introduction of the GMT is an ambitious reform and a major overhaul of the international corporate tax system. Under the GMT, large MNEs with revenues above €750 million are subject to a 15% minimum tax rate in each jurisdiction where they operate. Based on the Global Anti-Base Erosion (GloBE) Model Rules, the MNE's affiliate in the low-tax jurisdiction has to pay a top-up tax, if the jurisdictional effective tax rate (ETR) is below 15%. This can bring the total amount of tax on the MNE's excess profit – which is the GloBE income minus the Substance-Based Income Exclusion (SBIE) – in that jurisdiction up to the 15% rate. It is expected that the GMT will reduce MNEs' incentives of profit shifting and bring an end to the race to the bottom on corporate tax rates.

In this chapter, I study the effects of the GMT in a formal model of tax competition with two asymmetric countries that differ in market size. First, two countries choose corporate tax rates noncooperatively to maximize their revenues. Then the MNE chooses capital investment in each country and profit shifting level to maximize total after-tax profits, taking as given the tax environments. According to the OECD's Model Rules of December 2021, I assume that the top-up tax is collected by the host country where the MNE's affiliate records the profit. This corresponds to the qualified domestic minimum top-up tax (QDMTT). Following the previous literature on minimum taxation, I focus on the case where the minimum rate lies between two countries' initial tax rates without the GMT. As in Chapter 2 (and distinct from Chapter 1), countries' tax capacity – which relates to the cost of profit shifting – is assumed to be exogenous. But in contrast to Chapter 2, the (partial) deductibility rate is fixed in this chapter, and my focus is on the effects of the GMT rate and carve-out rate.

To the best of my knowledge, this chapter is the first to capture partial deductibility, the SBIE and QDMTT simultaneously in the international tax competition. I analyze how the GMT affects the MNE's behavior and countries' taxes from a short-run perspective where corporate tax rates are fixed, and from a long-run perspective where governments can adjust tax rates. Due to the SBIE, either increasing the minimum rate in the short run or setting tax rate below the minimum in the long run can exert investment incentives, while doing so also increases the top-up tax rate and incurs a larger revenue loss. This is the key trade-off in the analysis, which determines the short-run revenue effect of the GMT and reshapes the tax competition game in the long run.

The main results of Chapter 3 are as follows. Introducing the GMT can reduce the true tax differentials, curb profit shifting and always raise the large country's tax revenue. However, the revenue effect for the small country is ambiguous. In the short run where countries' tax rates are fixed, the GMT raises (reduces) the small country's tax revenue under high (low) profit shifting cost. In the long run where countries can adjust their tax rates, the GMT reshapes the underlying tax game and the competition pattern. At the equilibrium, the minimum rate binds the small country only if it is low. Otherwise, the small country will set its tax rate below the minimum, even at zero (if the carve-out rate is very small), in order to promote real investments. When the GMT rate is sufficiently high and the carve-out is not

very small, both countries undercut the minimum so that profit shifting vanishes. In this case, countries aim at attracting capital investments instead of competing for paper profits. Moreover, countries' equilibrium tax rates are decreasing in the minimum rate whenever they are below the minimum. While the equilibrium tax rate of the small country changes continuously with the minimum rate, gradual increases in the GMT rate may trigger the large country's tax rate to drop discontinuously below the minimum. As for the long-run revenue effect, a marginal GMT reform does not necessarily benefit the small country. It can raise the small country's tax revenue under large market-size asymmetry or low profit shifting cost, and may harm the small country otherwise. These results are substantially different from the traditional minimum tax literature which specifies the minimum tax as a lower bound imposed on countries' tax rates. Furthermore, I investigate the Pareto-improving non-marginal tax reform. When the carve-out rate is not too small, the GMT rate is not very high and profit shifting is not very sensitive to the minimum rate, the GMT can benefit the small country by increasing the taxation of both true profit and shifted profit.

Chapter 1. Tax Enforcement Leadership and Cooperation Against Profit

Shifting¹

Abstract

Market asymmetry between large and small countries induces tax gap that triggers profit shifting and base erosion from multinationals. Tax enforcement is the alternative to tax coordination to limit profit shifting. However, the lack of enforcement coordination makes the fight against profit shifting less effective. We consider a game in which countries differ in market size and enforcement productivity (enforcement elasticity of tax revenue). Countries seek to maximize welfare (tax revenue net of enforcement cost), choosing first their enforcement levels to limit profit shifting before competing in taxes. We find that enforcement leadership Pareto dominates simultaneous enforcement choices, and that the low-enforcement productivity country would be the leader. In line with the OECD/G20 BEPS project, we then analyze the scope for international enforcement cooperation. We establish that Nash bargaining over enforcements (with countries competing in taxes) induces higher enforcement, tax and revenue for each country, and that the benefit of enforcement cooperation is larger for the low-enforcement productivity country. We also show the robustness of our results to the endogenous tax leadership.

JEL classification: H25, H73, H87, C7

Keywords: Leadership, Tax enforcement, Profit shifting, Enforcement cooperation

¹ This chapter is co-authored with Prof. Jean Hindriks.

1.1. Introduction

With the prosperity of multinational enterprises (MNEs) in recent years, international tax issue on MNEs has taken center stage in public debates and policy circles. International tax-rate differentials provide MNEs with opportunities to shift profits between foreign subsidiaries in order to reduce their tax obligations. Overwhelming evidence shows that base erosion and profit shifting (BEPS) exist and cause considerable loss of tax revenue for both OECD and developing countries.² The impacts of BEPS on tax administrations worldwide are twofold. Firstly, induced by the profit shifting, countries compete with each other in corporation taxes so as to attract more mobile profits of MNEs, which results in substantial falls in corporate tax rates (see Devereux et al., 2008; Huizinga and Laeven, 2008; Overesch and Rincke, 2011).³ Secondly and more importantly, more and more countries realize the severity of BEPS and strengthen their tax enforcement to combat this problem⁴, which limits MNEs' profit shifting behavior (see Beer and Loeprick, 2015; Riedel et al., 2015; Dharmapala, 2019). Some recent empirical works further show that in addition to tax rates, countries may use tax enforcement as policy instrument to secure tax revenues, and find evidence of strategic interaction in tax enforcement among governments (e.g., Zinn et al., 2014; Durán-Cabré et al., 2015). In view of the above studies, Hindriks and Nishimura (2021) build a fiscal competition model where countries compete in tax enforcements simultaneously, and analyze the viability of enforcement cooperation. In their framework, however, the assumption of countries' simultaneous moves in choosing enforcement levels is questionable. Indeed, evidence suggests that countries' decisions on enhancing tax enforcement are not made simultaneously. For instance, as early as 2011, Chile changed the tax-enforcement policy on MNEs. It strongly increased reporting requirements for multinationals, increased the number of specialized tax auditors and created a specialized unit to monitor transfer pricing (see Bustos et al., 2019). Besides, in response to the OECD/G20 BEPS project, the progresses of enforcement reforms are also different among the participating countries. For instance, responding to interest deductions (Action 4) in the OECD/G20 BEPS package, Finland tightened the interest deductibility limits as early as 2013,

² For instance, Huizinga and Laeven (2008) points out that international profit shifting by multinationals causes Germany to lose significant tax revenues. Clausing (2016) finds that more than half of the foreign profits of US multinationals are booked in low-tax countries, which causes US tax revenue loss between \$77 billion and \$111 billion by 2012. Cristea and Nguyen (2016) shows that transfer price manipulation of physical good exports by multinationals reduces Danish government's corporate tax revenues by approximately 3.24%. Crivelli et al. (2016) argues that developing countries suffer from the problem of BEPS to a greater extent than OECD countries. Bilicka (2019) reveals that foreign multinational subsidiaries in the UK underreport taxable profits by about 50 percent relative to (comparable) domestic companies. See also OECD's (2015) final reports.

³ In 1980, corporate tax rates around the world averaged 40.11 percent, and 46.52 percent when weighted by GDP. In 2020, the average is 23.85 percent, and 25.85 when weighted by GDP (see Tax Foundation, "Corporate Tax Rates around the World, 2020," Dec. 9, 2020, <https://taxfoundation.org/corporate-tax-rates-around-the-world-2020/>).

⁴ In particular, G20 and OECD launch a BEPS package which provides 15 action plans against tax avoidance by multinationals (see OECD, 2015). OECD and G20 countries along with an increasing number of developing countries actively participate in the development of this OECD/G20 BEPS project, and make unparalleled efforts to tackle BEPS. The measures that countries often take to enhance tax enforcement include designing effective Controlled Foreign Company (CFC) rules, tightening transfer pricing regulations and increasing the auditing probability (see e.g., Buettner and Wamser, 2013; Durán-Cabré et al., 2015; Riedel et al., 2015).

while a draft Financial Bill including similar rules was proposed in France until 2018. Regarding transfer pricing documentation (Action 13), a law introducing the country-by-country (CbC) reporting requirement took effect in Italy in January, 2016. In contrast, Czech Republic implemented the CbC reporting requirements nearly two years later.⁵ These facts naturally raise the following questions: how does the sequence of enforcement decisions affect countries' tax revenues? Which countries are likely to lead in tax enforcement competition if the timing of enforcement is taken into account, and what is the determinant of the enforcement leadership? To the best of our knowledge, there is no study on the tax enforcement leadership.

In this paper, we build upon Hindriks and Nishimura's (2021) model by endogenizing the timing of countries' enforcement decisions. To this end, we introduce a timing game at the pre-play stage, as proposed by Hamilton and Slutsky (1990), in which two countries commit themselves to move early or late in choosing tax enforcements.⁶ At the next stage, two countries choose non-cooperatively their enforcement levels, based on their timing decisions at the pre-play stage. Through this approach, we stipulate an "enforcement leadership game" to analyze the emergence of the enforcement leadership. Another contribution of this paper is to consider asymmetric enforcement productivity among countries.⁷ Using a Cobb-Douglas enforcement technology, we represent the enforcement productivity as the (constant) enforcement elasticity of tax revenue.⁸ This is related to Keen and Slemrod (2017), who propose the concept "enforcement elasticity of tax revenue" to capture the productivity of the administrative interventions (i.e., enforcement). Throughout the paper, we refer to the country that has a high (low) enforcement elasticity of tax revenue as the high (low) enforcement productivity country.⁹ Given the super-modularity of the Cobb-Douglas function, following Amir and Stepanova (2006), we find that the subgame perfect equilibria (SPEs) of the enforcement leadership game consist of two Stackelberg equilibria in which countries set their enforcement levels sequentially. Both countries are better off under Stackelberg equilibria than under simultaneous Nash equilibrium. We further show that the leadership by the low-enforcement productivity country risk dominates. Intuitively, by being an enforcement leader, the low-enforcement productivity country can

⁵ The information on countries' action developments in response to the OECD/G20 BEPS project is compiled from Ernst & Young's survey (more details available at: https://www.ey.com/en_gl/beps-tracker).

⁶ The game with endogenous timing of moves, which is also called "commitment game", has been well developed in tax competition model to analyze the leadership of tax decisions. See, for instance, Kempf and Rota-Graziosi (2010), Hindriks and Nishimura (2015, 2017) for the analysis of leadership among countries in setting tax rates. In our base model, the tax leadership is absent: we exogenously assume that two countries set tax rates simultaneously so as to focus on the enforcement leadership. However, as demonstrated in Section 1.4, incorporating the tax leadership into our framework does not change the main results obtained in the base model.

⁷ Various forms of country asymmetries have been studied in the fiscal competition literature, including the asymmetries in endowments, market size and capital ownership (see Bucovetsky, 1991; Peralta and van Ypersele, 2005; Hindriks and Nishimura, 2015, 2017, 2021).

⁸ Several concepts concerning tax evasion issues have been developed in the previous literature, e.g., concealment technology, tax avoidance technology and tax collection technology. See Slemrod and Yitzhaki (2002) for a comprehensive review of this field. In our paper, the enforcement technology transforms countries' enforcement efforts into the cost of profit shifting for multinationals.

⁹ Hindriks and Nishimura (2021) assume that the enforcement technology is a symmetric CES function with the symmetric Cobb-Douglas as a special case. The asymmetry of enforcement productivity is therefore ruled out in their model. We discuss further the specification of the enforcement technology in Subsection 1.2.3.

gain more revenue from the follower’s enforcement response, than the other way round. This triggers it to take the lead in the enforcement competition.

Although both countries are better off by choosing enforcements sequentially, we show that there is still scope for mutually beneficial cooperation on tax enforcements. Unlike Hindriks and Nishimura (2021), in which cooperation maximizes joint welfare (with the risk that the low-tax country may lose from such cooperation in the absence of compensating transfers), we study international enforcement cooperation as a Nash bargaining problem.¹⁰ Some recent literature on profit shifting also involves Nash bargaining, such as bargaining between the MNE and labor union, between the headquarter and supplier (Riedel, 2011; Bauer and Langenmayr, 2013). Becker and Davies (2014) provide a “negotiation model” in which the high-tax country and low-tax country bargain over the multinational’s transfer price, but tax competition is absent in their analysis. In contrast, our paper analyzes bargaining over enforcement levels with countries competing in tax rates. We present that bargaining outcome leads to higher enforcements, taxes and welfare for each country relative to the Stackelberg outcomes. Furthermore, the welfare gain (in percentage) from enforcement cooperation is larger for the low-enforcement productivity country. The rationale is that the positive enforcement externality is larger on the country with a lower enforcement productivity (see Subsection 1.3.2). Consequently, it can benefit more from the internalization of enforcement externality under Nash bargaining. Our results are also shown to be robust to the sequential choices of taxes (tax leadership). Compared to the simultaneous tax competition, tax leadership by the large country raises two countries’ equilibrium enforcements, taxes and welfare under both enforcement leadership and Nash bargaining.

Our paper highlights the key role of enforcement productivity. First, the asymmetric enforcement productivity drives the enforcement leadership, which is different from the driving force of tax leadership (i.e., the asymmetric market size) identified in the tax competition literature. The country with low enforcement productivity will be the enforcement leader. Second, the low-enforcement productivity country benefits (relatively) more from the enforcement cooperation. Third, the welfare effect of the market-size asymmetry is affected by the enforcement productivity. In particular, increasing market-size disparity can harm the large country if its enforcement productivity is sufficiently low.

The paper is organized as follows. Section 1.2 presents the model and the non-cooperative equilibrium with endogenous enforcement leadership. Section 1.3 investigates enforcement cooperation under Nash bargaining. Section 1.4 extends the base model to analyze tax leadership. Section 1.5 concludes. All the proofs are in Appendix 1.A.

1.2. The model

¹⁰ Hindriks and Nishimura (2021) showed that such cooperation may not benefit the low-tax country if the market-size asymmetry is too large. In contrast, the Nash bargaining outcome will necessarily make both countries better off compared to the non-cooperative outcome.

1.2.1. The setups

Building upon Hindriks and Nishimura's (2021) model, we consider two countries indexed by 1 and 2 and two MNEs indexed by a and b . Each MNE has affiliates in both countries, and produces a homogeneous good with unit cost c_i in country i ($i=1,2$). The output of MNE k ($k=a,b$) in country i is denoted by q_i^k , and the inverse demand function in country i is $p_i = \alpha_i - \beta(q_i^a + q_i^b)$.

Without loss of generality, we assume $\gamma_1 := \alpha_1 - c_1 > \gamma_2 := \alpha_2 - c_2$. This assumption introduces the asymmetry in market size: country 1 with greater market demand (or lower supply cost) is the large country, while country 2 is the small country. Following Hindriks and Nishimura (2021), we define $\varepsilon := \frac{1 - (\gamma_2 / \gamma_1)^2}{1 + (\gamma_2 / \gamma_1)^2} \in (0,1]$ as the index of market size asymmetry.

The profit of MNE k ($k=a,b$) in country i is $\pi_i^k = (p_i - c_i)q_i^k$. Each MNE chooses the amount of profit to be shifted, so as to minimize its total tax liabilities net of the profit shifting cost. The profit shifting (concealment) cost is:

$$C(\pi_i^k, \tilde{\pi}_i^k) = \delta(e_1, e_2)(\pi_1^k - \tilde{\pi}_1^k)^2 + \delta(e_1, e_2)(\pi_2^k - \tilde{\pi}_2^k)^2 = 2\delta(e_1, e_2)(\pi_1^k - \tilde{\pi}_1^k)^2,$$

where π_i^k , $\tilde{\pi}_i^k$ and e_i respectively represent the MNE k 's actual profit, reported profit and the level of enforcement in country i .

The MNE k 's overall after-tax profit is $\sum_i (1-t_i)\tilde{\pi}_i^k - C(\pi_i^k, \tilde{\pi}_i^k)$, where t_i denotes the source-based corporate tax rate in country i .

$\delta(e_1, e_2)$ denotes the enforcement technology and represents the overall enforcement level, which transforms two countries' enforcement efforts into the cost associated with MNEs' profit shifting. We assume that $\delta(e_1, e_2)$ is strictly increasing and concave in each

country's enforcement level, i.e., $\forall e_1, e_2 > 0$, $\frac{\partial \delta(e_i, e_j)}{\partial e_i} > 0$, $\frac{\partial^2 \delta(e_i, e_j)}{\partial e_i^2} \leq 0$. When one

country raises its enforcement (such as setting tighter transfer pricing laws or auditing MNEs more frequently), the cost of (successful) profit shifting increases for the MNEs.

Tax enforcement is costly, as countries have to devote real resources to expand enforcement activities. We denote by $EC_i(e_i)$ the enforcement cost of country i with $EC'_i(e_i) > 0$ and $EC''_i(e_i) > 0$, $\forall e_i > 0$.¹¹ Then the welfare (i.e., the tax revenue net of the enforcement cost) of country i is:¹²

$$W_i = t_i(\tilde{\pi}_i^a + \tilde{\pi}_i^b) - EC_i(e_i). \quad (1.1)$$

1.2.2. Enforcement leadership

To endogenize the enforcement leadership, we introduce a timing game at stage 1 (i.e., the pre-play stage). The sequence of players' decisions is as below:

Stage 1 (pre-play stage): Countries non-cooperatively decide whether to move “Early” or “Late” in choosing their enforcement levels.

Stage 2: Countries non-cooperatively choose enforcement levels to maximize their welfare, according to the timing decisions at stage 1.

Stage 3: Countries set simultaneously and non-cooperatively their tax rates to maximize their tax revenues.

Stage 4: MNEs choose outputs and reported profits in each country to maximize their after-tax profits.

There are four strategy pairs at the pre-play stage: (Early, Early), (Early, Late), (Late, Early), (Late, Late), which can generate three different games at stage 2. Specifically, a simultaneous-move game G^N is induced when two countries make the same choice (i.e., (Early, Early) or (Late, Late)). When country i moves early and the other moves late, it induces a sequential game G^i , where superscript i represents the enforcement leadership by country i .

¹¹ The enforcement cost, which is also referred to as the administrative cost of enforcement, may result from employing tax inspectors and from the burden of examining statutory documentation provided by multinationals (see Zinn et al., 2014). Almunia and Lopez-Rodriguez (2018) also specify a strictly increasing and convex function to capture the enforcement cost when analyzing the effects of tax enforcement on firms' tax compliance and on social welfare. See also Slemrod and Yitzhaki (2002), who discussed at length the administrative cost from the perspective of information gathering.

¹² As in Hindriks and Nishimura (2021), the profit tax with full deductibility does not distort the productions and prices, which means that the consumer surplus is independent of the tax system. Therefore, we refer to the tax revenue net of the enforcement cost as the country's welfare throughout the paper.

Hereafter, the leader and follower in each sequential game are respectively denoted by superscripts L and F , while in the simultaneous Nash game G^N two countries are denoted by superscript N .

1.2.3. The equilibrium analysis

The backward induction is applied to solve for the subgame perfect equilibria (SPEs) of the game. Let $\beta = \frac{2}{9}(\gamma_1^2 + \gamma_2^2)$ such that the total profits of the two MNEs are normalized to unity. Let t_i^* and R_i^* be respectively country i 's equilibrium tax rate and revenue in the stage-3 tax subgame. Lemma 1 gives the equilibrium outcomes of stages 3 and 4 (see Equations (4), (7) and (9) in Hindriks and Nishimura (2021)).

Lemma 1. (i) $\pi_i^k = \frac{1+\mathbf{1}\varepsilon}{4}$, $\tilde{\pi}_i^k = \frac{1+\mathbf{1}\varepsilon}{4} - \frac{t_i - t_j}{4\delta(e_1, e_2)}$ ($k = a, b$).

$$(ii) \quad t_i^* = \frac{3+\mathbf{1}\varepsilon}{3}\delta(e_1, e_2), \quad R_i^* = \frac{(3+\mathbf{1}\varepsilon)^2}{18}\delta(e_1, e_2),$$

where $\mathbf{1}$ is an indicator taking the value 1 for country 1 and -1 for country 2.

Lemma 1(ii) shows that the large country (i.e., country 1) sets a higher tax than its rival in the (simultaneous) tax subgame. The equilibrium tax gap is driven by the market-size asymmetry. To see this more clearly, consider each country's tax base elasticity under uniform rate:

$$-\frac{\partial \tilde{\pi}_1}{\partial t_1} \frac{t_1}{\tilde{\pi}_1} = \frac{t}{(1+\varepsilon)\delta(e_1, e_2)} < -\frac{\partial \tilde{\pi}_2}{\partial t_2} \frac{t_2}{\tilde{\pi}_2} = \frac{t}{(1-\varepsilon)\delta(e_1, e_2)}, \quad \text{where } \tilde{\pi}_i = \sum_k \tilde{\pi}_i^k. \quad \text{So, the small}$$

country (i.e., country 2) perceives its tax base to be more elastic and thus taxes less. Note that tax base elasticities are inversely proportional to the overall enforcement level. So, the benefit of enforcement is to enhance the fiscal capacity of each country by limiting their tax base elasticity. Interestingly, the low tax country also benefits from enforcement, even though for any given tax rates, the amount of profit shifting is decreasing with enforcement. The reason is that in response to higher enforcement, the low tax country will increase its tax to a lesser extent than the high tax country (Lemma 1(ii)), so as to offset the impact of enforcement on profit shifting. Indeed, from Lemma 1, profit shifting is, under equilibrium taxes, independent of the overall enforcement. The implication is that both countries can benefit from higher

enforcement but face the free rider incentive.

Using Lemma 1 and (1.1), the welfare function at stage 2 is:

$$W_i(e_1, e_2) = d_i \delta(e_1, e_2) - EC_i(e_i), \quad (1.2)$$

where $d_i = \frac{(3+\mathbf{1}\varepsilon)^2}{18}$.

Following Amir and Stepanova (2006), the stage 1 equilibria depend on the cross partial derivative of $\delta(e_1, e_2)$. Specifically, if $\frac{\partial^2 \delta(e_1, e_2)}{\partial e_1 \partial e_2} > 0$, $\forall e_1, e_2 > 0$, then both (Early, Late) and (Late, Early) are the equilibria of the enforcement timing game and can lead to higher welfare for both countries than the simultaneous game. If $\frac{\partial^2 \delta(e_1, e_2)}{\partial e_1 \partial e_2} < 0$, $\forall e_1, e_2 > 0$, then (Early, Early) is the unique equilibrium of the enforcement timing game. Two countries are indifferent between moving early and late, and all the four strategy pairs are the equilibria if $\frac{\partial^2 \delta(e_1, e_2)}{\partial e_1 \partial e_2} \equiv 0$, $\forall e_1, e_2 > 0$.¹³

To further analyze the enforcement leadership, we impose some structure on the administrative cost and enforcement technology by specifying a quadratic enforcement cost function $EC_i(e_i) := \frac{\eta}{2} e_i^2$ ($\eta > 0$) and a Cobb-Douglas enforcement technology $\delta(e_1, e_2) := e_1^\alpha e_2^{1-\alpha}$ ($0 < \alpha < 1$).¹⁴ Given (2) and the above specifications, country i 's welfare

¹³ Following Amir and Stepanova's (2006) approach, given the welfare function (2), we can establish that for each country: (i) the equilibrium welfare in the sequential game where it follows is higher (resp. lower) than in game G^N if $\frac{\partial^2 \delta(e_1, e_2)}{\partial e_1 \partial e_2} > 0$, $\forall e_1, e_2 > 0$ (resp. $\frac{\partial^2 \delta(e_1, e_2)}{\partial e_1 \partial e_2} < 0$, $\forall e_1, e_2 > 0$); (ii) the equilibrium welfare in the sequential game where it leads is always higher than in game G^N if $\frac{\partial^2 \delta(e_1, e_2)}{\partial e_1 \partial e_2} \neq 0$, $\forall e_1, e_2 > 0$; and (iii) its equilibrium enforcement and welfare levels are the same in games G^1 , G^2 and G^N if $\frac{\partial^2 \delta(e_1, e_2)}{\partial e_1 \partial e_2} \equiv 0$, $\forall e_1, e_2 > 0$. These properties directly yield the stage-1 equilibria. The detailed derivations are available upon request.

¹⁴ Empirical evidence indicates that countries strengthen their transfer pricing regulations in response to others' tighter policies against profit shifting (see Zinn et al., 2014; Riedel et al., 2015). This supports the specification of a super-modular enforcement technology with $\frac{\partial^2 (e_1^\alpha e_2^{1-\alpha})}{\partial e_1 \partial e_2} > 0$, $\forall e_1, e_2 > 0$, since in our model strategic complementarity (resp. substitutability) of tax enforcements corresponds to the super-modularity (resp. sub-modularity) of $\delta(e_1, e_2)$. The related

function at stage 2 is:

$$W_i(e_1, e_2) = d_i e_1^\alpha e_2^{1-\alpha} - \frac{\eta}{2} e_i^2. \quad (1.3)$$

The Cobb-Douglas enforcement technology implies constant enforcement elasticity of tax revenue in each country, based on which we can introduce the asymmetry in countries' enforcement productivity.¹⁵ Specifically, parameters α and $1-\alpha$ respectively represent the enforcement elasticity of tax revenue in country 1 and in country 2. In line with Keen and Slemrod (2017), we say that enforcement productivity is higher (resp. lower) in country 1 than in country 2 if $\alpha > \frac{1}{2}$ (resp. $\alpha < \frac{1}{2}$).¹⁶ Hence, our model involves two types of asymmetry: the asymmetry in market size and the asymmetry in enforcement productivity.¹⁷

To secure interior solutions in the tax subgame, we assume that the enforcement cost parameter η is sufficiently large with $\eta \geq \frac{8}{3}$. Given (1.3), the equilibrium enforcement and welfare levels associated with games G^N , G^1 and G^2 are displayed in Table 1.1:

derivations are available upon request.

¹⁵ Note that under the Cobb-Douglas enforcement technology, the alternative timing in which countries simultaneously choose enforcements and taxes would lead to trivial solutions, i.e., $\tilde{e}_1 = \tilde{e}_2 = 0$ and $\tilde{t}_1 = \tilde{t}_2 = 0$ (detailed derivations are available on request). The reason is that with Cobb-Douglas technology, both countries' participation in tax enforcement is required to secure effective enforcement. Each country has an incentive to tax less than the other and to set its enforcement to zero, so as to capture all the profits of the MNEs. This extreme "race to the bottom" outcome obviously does not fit the reality. Moreover, enforcement levels, usually determined by monitoring policies and transfer pricing rules, are long-term policy variables and are altered far less frequently than corporate tax rates. Therefore, it is reasonable to assume that countries pre-commit on enforcements before setting their taxes. See Peralta et al. (2006), Bucovetsky and Haufler (2008), and Hindriks and Nishimura (2021) for similar setups.

¹⁶ Suppose that each country unilaterally increases its enforcement level by one percent. Then country 1 can raise its revenue by a larger percentage than country 2 iff $\alpha > \frac{1}{2}$.

¹⁷ For simplicity, throughout the paper we focus on the Cobb-Douglas function that is homogeneous of degree 1, such that the sum of enforcement productivities is normalized to unity. All the results in our paper remain qualitatively unchanged under a more general Cobb-Douglas enforcement technology: $\delta(e_1, e_2) := e_1^\alpha e_2^\beta$ with $\alpha, \beta > 0$ and $0 < \alpha + \beta \leq 1$. Under this general form, country 1 displays a higher enforcement productivity iff $\alpha > \beta$.

Table 1.1: Equilibrium enforcement and welfare in games G^N , G^1 and G^2 .

	G^N	G^1	G^2
Country 1			
e_1	$\frac{(1-\alpha)^{\frac{1-\alpha}{2}} \alpha^{\frac{1+\alpha}{2}} d_1^{\frac{1+\alpha}{2}} d_2^{\frac{1-\alpha}{2}}}{\eta}$	$\frac{\left(\frac{2\alpha}{1+\alpha}\right)^{\frac{1+\alpha}{2}} (1-\alpha)^{\frac{1-\alpha}{2}} d_1^{\frac{1+\alpha}{2}} d_2^{\frac{1-\alpha}{2}}}{\eta}$	$\frac{\left(\frac{2-2\alpha}{2-\alpha}\right)^{\frac{1-\alpha}{2}} \alpha^{\frac{1+\alpha}{2}} d_1^{\frac{1+\alpha}{2}} d_2^{\frac{1-\alpha}{2}}}{\eta}$
W_1	$\frac{\alpha^\alpha (2-\alpha)(1-\alpha)^{1-\alpha} d_1^{1+\alpha} d_2^{1-\alpha}}{2\eta}$	$\frac{2^\alpha \alpha^\alpha (1-\alpha)^{1-\alpha} (1+\alpha)^{-1-\alpha} d_1^{1+\alpha} d_2^{1-\alpha}}{\eta}$	$\frac{2^{-\alpha} \alpha^\alpha (2-\alpha)^\alpha (1-\alpha)^{1-\alpha} d_1^{1+\alpha} d_2^{1-\alpha}}{\eta}$
Country 2			
e_2	$\frac{(1-\alpha)^{\frac{2-\alpha}{2}} \alpha^{\frac{\alpha}{2}} d_1^{\frac{\alpha}{2}} d_2^{\frac{2-\alpha}{2}}}{\eta}$	$\frac{\left(\frac{2\alpha}{1+\alpha}\right)^{\frac{\alpha}{2}} (1-\alpha)^{\frac{2-\alpha}{2}} d_1^{\frac{\alpha}{2}} d_2^{\frac{2-\alpha}{2}}}{\eta}$	$\frac{\left(\frac{2-2\alpha}{2-\alpha}\right)^{\frac{2-\alpha}{2}} \alpha^{\frac{\alpha}{2}} d_1^{\frac{\alpha}{2}} d_2^{\frac{2-\alpha}{2}}}{\eta}$
W_2	$\frac{(1+\alpha)\alpha^\alpha (1-\alpha)^{1-\alpha} d_1^\alpha d_2^{2-\alpha}}{2\eta}$	$\frac{2^{\alpha-1} \alpha^\alpha (1-\alpha)^{1-\alpha} (1+\alpha)^{1-\alpha} d_1^\alpha d_2^{2-\alpha}}{\eta}$	$\frac{2^{1-\alpha} \alpha^\alpha (2-\alpha)^{\alpha-2} (1-\alpha)^{1-\alpha} d_1^\alpha d_2^{2-\alpha}}{\eta}$

From Table 1.1, we can compare the equilibrium enforcement and welfare levels under three games:

$$\forall \alpha \in (0,1), \quad e_i^L > e_i^F > e_i^N \quad \text{and} \quad W_i^F > W_i^L > W_i^N, \quad (1.4)$$

where $W_i^L := W_i(e_i^L, e_j^F)$, $W_i^F := W_i(e_i^F, e_j^L)$ and $W_i^N := W_i(e_i^N, e_j^N)$ respectively represent the equilibrium payoffs of country i in games G^i , G^j and G^N ($i, j=1,2$; $i \neq j$).

Compared to the Nash equilibrium, Stackelberg equilibria induce higher enforcement levels in both countries. Indeed, consider a small increase in the leader's enforcement relative to the Nash equilibrium. By strategic complementarity, this raises the follower's enforcement, which in turn benefits the leader because of the positive enforcement externality (higher enforcement reduces the tax base elasticity of two countries, enabling them to tax more).

Given the welfare ranking in (1.4), each country prefers to follow than to lead (second-mover advantage), but at the same time both countries prefer sequential choice to simultaneous move. As a result, the two Stackelberg equilibria in which each country can either lead or follow constitute the SPEs of the timing game.

To solve the coordination problem caused by multiple equilibria, we use the Pareto-dominance and risk dominance refinement (see Harsanyi and Selten, 1988).¹⁸ The

¹⁸ Even if neither of the two SPEs is Pareto-dominant (i.e., Proposition 1(ii)), the risk-dominant equilibrium is often selected in the coordination game. This is because it is perceived as a safer choice when players face strategic uncertainty about rivals'

equilibrium (Early, Late) risk dominates (Late, Early) iff:

$$\Pi := (W_1^L - W_1^N)(W_2^F - W_2^N) - (W_1^F - W_1^N)(W_2^L - W_2^N) > 0.$$

Then, we can establish the following:

Proposition 1. (i) *The enforcement leadership by country 1 (resp. country 2) is risk-dominant iff $0 < \alpha < \frac{1}{2}$ (resp. $\frac{1}{2} < \alpha < 1$). If $\alpha = \frac{1}{2}$, there is no risk-dominant enforcement leadership.*
(ii) *Neither of the Stackelberg equilibria Pareto-dominates the other.*

Proof. See Appendix 1.A.1.

This proposition contrasts with the tax leadership derived from the market-size asymmetry (see Kempf and Rota-Graziosi, 2010; Hindriks and Nishimura, 2015, 2017). In our model, the enforcement leadership is driven by the enforcement strategic complementarity together with the positive externality of enforcement. Indeed, from lemma 1 (ii), we see that even the low-tax country benefits from higher enforcements, because in equilibrium it will increase its tax to a lesser extent than the high-tax country so as to maintain its tax base unchanged. This is different from usual analysis of tax enforcement in which taxes are fixed. In our model, enforcement enables countries to scale up their tax. The enforcement leadership is taken by the low-enforcement productivity country. The reason is that the positive externality induced by the follower's enforcement response is larger when the follower has a higher enforcement productivity. To see this, we present country i 's ($i = 1, 2$) reaction function $e_i^N(e_j) := \arg \max_{e_i} W_i(e_i, e_j)$ in game G^N :

$$\begin{cases} e_1^N(e_2) = \left(\frac{\alpha d_1}{\eta}\right)^{\frac{1}{2-\alpha}} e_2^{\frac{1-\alpha}{2-\alpha}}, \\ e_2^N(e_1) = \left[\frac{(1-\alpha)d_2}{\eta}\right]^{\frac{1}{1+\alpha}} e_1^{\frac{\alpha}{1+\alpha}}. \end{cases} \quad (1.5)$$

As footnote (14) states, the super-modularity of the enforcement technology implies that enforcements are strategic complements: $\frac{\partial e_i^N(e_j)}{\partial e_j} > 0$.

In game G^i , denote by $e_j^F(e_i)$ follower j 's best response function. Then, leader i 's

moves (see Van Huyck et al, 1990; Amir and Stepanova, 2006).

tax revenue function is $R_i(e_i, e_j^F(e_i)) = d_i \delta(e_i, e_j^F(e_i))$, from which we can derive the leader's enforcement elasticity of tax revenue in game G^i :¹⁹

$$\frac{dR_i / R_i}{de_i / e_i} = \underbrace{\frac{\partial R_i / R_i}{\partial e_i / e_i}}_{\text{enforcement productivity}} + \underbrace{\frac{\partial R_i / R_i}{\partial e_j^F / e_j} \cdot \frac{de_j^F / e_j}{de_i / e_i}}_{\text{interaction effect}}. \quad (1.6)$$

Suppose that country i in game G^i increases its enforcement level by one percent from the Nash level e_i^N . This triggers country j to raise its enforcement by $\frac{de_j^F / e_j}{de_i / e_i}$ percent. From (1.6), the strategic response of country j can raise country i 's revenue by

$\frac{\partial R_i / R_i}{\partial e_j^F / e_j} \cdot \frac{de_j^F / e_j}{de_i / e_i}$ percent. We refer to it as the “interaction effect” on leader i 's revenue in

game G^i . As shown by (1.6), the effect of the leader's enforcement on its own revenue in the Stackelberg game can be decomposed into two parts. The first part is its enforcement productivity, and the second part is the interaction effect, which measures the enforcement externality from the follower's strategic response.

Notice that follower j is on its best response function in Stackelberg game G^i , which means that $e_j^F(e_i) \equiv e_j^N(e_i)$. Then given (1.5) and (1.6), the interaction effects in games G^1 and G^2 are $\frac{\alpha(1-\alpha)}{1+\alpha}$ and $\frac{\alpha(1-\alpha)}{2-\alpha}$, respectively. Comparing the two values, we see that the interaction effect in game G^1 is larger (resp. smaller) than in game G^2 if $0 < \alpha < \frac{1}{2}$ (resp. $\frac{1}{2} < \alpha < 1$). Therefore, when $0 < \alpha < \frac{1}{2}$, country 1 gains more as a leader from country 2's enforcement response, than the other way round. In other words, country 1 loses more than country 2 by reverting to simultaneous moves instead of being the enforcement leader. This triggers country 1 (the low enforcement productivity country) to take the lead in the enforcement competition. The opposite reasoning holds for $\frac{1}{2} < \alpha < 1$, where a switch of the relative enforcement productivity induces a switch of enforcement leadership.

¹⁹ The strategic complementarity of tax enforcements produces a gap between the exogenous enforcement productivity and the endogenous enforcement elasticity of tax revenue under enforcement leadership.

1.3. Cooperative enforcement

1.3.1. Nash bargaining over enforcement

Thus far, we assume that enforcement levels are set non-cooperatively, and our findings reveal that both countries are better off by setting enforcements sequentially instead of simultaneously. Nevertheless, there is still scope for efficiency gains from enforcement coordination, as indicated by Lemma 2:

Lemma 2. *For any $\alpha \in (0,1)$, a small equi-proportional increase of both countries' enforcements from the Stackelberg levels is strictly Pareto-improving.*

Proof. See Appendix 1.A.2.

This lemma points out the inefficiency of the two Stackelberg equilibria, as both countries can be better off by slightly increasing their enforcements from the Stackelberg levels. As mentioned before, the enforcement leadership increases both countries' enforcement levels compared to the simultaneous Nash game. Nevertheless, the positive externality of enforcement cannot be fully internalized in the Stackelberg games. When choosing its enforcement, each country does not take into account the positive effect of its enforcement on the other (to enhance its tax capacity). As a result, the Stackelberg equilibria involve suboptimal enforcement levels.

This inefficiency provides scope for international enforcement cooperation. In contrast with the marginal enforcement reform indicated by Lemma 2 (Pareto improvement), a more practical policy is to negotiate over enforcement to achieve the best possible outcome (Pareto efficiency). In the following, we consider enforcement cooperation as a Nash bargaining problem. Specifically, two countries negotiate over enforcement levels to maximize the product of their welfare gains relative to the disagreement outcomes. In case the negotiation breaks down, they choose enforcement levels according to the enforcement leadership game.²⁰

²⁰ Hindriks and Nishimura (2021) treat enforcement cooperation in a different way. They consider enforcement cooperation as maximization of the joint welfare. In the absence of (lump sum) compensating transfers, enforcement cooperation is beneficial for both countries only if the market-size asymmetry is small enough, otherwise the small (low-tax) country will lose from the cooperation. By contrast, Nash bargaining solution rules out the possible disagreement against cooperation,

Formally, based on Proposition 1, the Nash bargaining problem to be solved is:²¹

$$\max_{\{e_1, e_2\}} (W_1(e_1, e_2) - W_1^u)(W_2(e_1, e_2) - W_2^v), \quad (1.7)$$

subject to :

$$e_1 \geq 0, \quad e_2 \geq 0, \quad W_1(e_1, e_2) \geq W_1^u \quad \text{and} \quad W_2(e_1, e_2) \geq W_2^v,$$

$$\text{where } (u, v) = \begin{cases} (L, F) & \text{if } \alpha \in (0, \frac{1}{2}) \\ (F, L) & \text{if } \alpha \in (\frac{1}{2}, 1) \end{cases}.$$

The strict concavity of the welfare functions guarantees that the feasible set is convex and the Nash product is strictly quasi-concave, which together with the compact feasible set ensure the existence and uniqueness of the solution to Problem (1.7).²² Hereafter, we use superscript B to denote the Nash bargaining solution for each country.

1.3.2. Bargaining solution

We can compare the equilibrium outcomes under Nash bargaining and enforcement leadership game.

Proposition 2. *The equilibrium enforcement levels, tax rates, and welfare levels of both countries are strictly higher under Nash bargaining than in the (non-cooperative) enforcement leadership game.*

Proof. See Appendix 1.A.3.

Although both countries can enhance enforcement levels by moving sequentially instead of simultaneously, their enforcements can be further increased to efficient levels under Nash Bargaining. The fundamental mechanism behind this result is twofold. First, the positive

since both countries are necessarily better off at this solution compared to non-cooperation. Another difference from Hindriks and Nishimura (2021) is that the disagreement outcome in our model is the Stackelberg equilibrium and not the simultaneous Nash equilibrium.

²¹ Here, two countries are supposed to possess equal bargaining power. The assumption is reasonable. Indeed, each country can still secure Stackelberg outcomes without participating in enforcement cooperation. Neither country will have a superior bargaining position to the other in the negotiation.

²² The detailed proofs are available on request to the authors.

enforcement externality $\frac{\partial W_i(e_i, e_j)}{\partial e_j} > 0$ ($i, j = 1, 2 ; i \neq j$) is internalized under Nash

bargaining, since both countries are maximizing the product of their welfare gain (relative to the disagreement outcomes) instead of their own welfare. When bargaining over the enforcement, each country takes into account the positive externality of its enforcement on the other country. Thus, both countries have incentives to exert more enforcement efforts, which enables them to raise taxes (without changing their tax base). Second, due to the

super-modularity of the enforcement technology $\frac{\partial^2 \delta(e_i, e_j)}{\partial e_i \partial e_j} > 0$ ($i, j = 1, 2 ; i \neq j$), higher

enforcement by one country increases the marginal welfare of enforcement for the other country. Therefore, the other country has an incentive to enhance enforcement, too. The two forces act in the same direction to raise enforcement levels compared to the Stackelberg levels.²³ The increased enforcements limit profit shifting, reduce the tax base elasticity and induce higher taxes under Nash bargaining.

We now analyze the cooperative welfare under Nash bargaining. The following proposition states which country gains (relatively) more from enforcement cooperation.

Proposition 3. *Compared to (Stackelberg) disagreement outcomes, under enforcement cooperation the leader's welfare gain (in percentage) is larger than the follower's welfare gain, i.e.,*

$$\left\{ \begin{array}{l} \frac{W_1^B}{W_1^L} > \frac{W_2^B}{W_2^F} \text{ if } \alpha \in (0, \frac{1}{2}) \\ \frac{W_2^B}{W_2^L} > \frac{W_1^B}{W_1^F} \text{ if } \alpha \in (\frac{1}{2}, 1) \end{array} \right.$$

Proof. See Appendix 1.A.4.

Although two countries have the same bargaining power, their welfare levels do not increase in the same proportion. The explanation is as follows. Enforcement cooperation leads both countries to raise enforcements above the Stackelberg levels. Recall that each country's

²³ Proposition 2 always holds for a super-modular enforcement technology, i.e., $\forall e_1, e_2 > 0, \frac{\partial^2 \delta(e_i, e_j)}{\partial e_i \partial e_j} > 0$ ($i = 1, 2 ; i \neq j$). See Appendix 1.A.3 for a proof without specifying the functional form of $\delta(e_1, e_2)$.

revenue function is $R_i = d_i e_1^\alpha e_2^{1-\alpha}$, which yields $\frac{\partial R_1 / R_1}{\partial e_2 / e_2} = 1 - \alpha$ and $\frac{\partial R_2 / R_2}{\partial e_1 / e_1} = \alpha$. Hence, the

positive enforcement externality (in percentage) on one country is equal to the enforcement productivity of the other country. This means that the enforcement externality is larger on the country with a lower enforcement productivity. Therefore, relative to the Stackelberg outcome, the low-enforcement productivity country (i.e., the enforcement leader) benefits more from enforcement cooperation.

Proposition 3 also implies that the distribution of the efficiency gains from cooperation is twisted in favor of the enforcement leader. To see this, we rewrite the expression in Proposition 3 as below:

$$\left\{ \begin{array}{l} \frac{W_1^B}{W_1^B + W_2^B} > \frac{W_1^L}{W_1^L + W_2^F} \\ \frac{W_2^B}{W_1^B + W_2^B} < \frac{W_2^F}{W_1^L + W_2^F} \end{array} \right. \text{ if } 0 < \alpha < \frac{1}{2}, \text{ and } \left\{ \begin{array}{l} \frac{W_1^B}{W_1^B + W_2^B} < \frac{W_1^F}{W_1^F + W_2^L} \\ \frac{W_2^B}{W_1^B + W_2^B} > \frac{W_2^L}{W_1^F + W_2^L} \end{array} \right. \text{ if } \frac{1}{2} < \alpha < 1.$$

This shows that compared to the Stackelberg equilibrium, cooperation increases the enforcement leader's share in the total welfare. In summary, the enforcement leader benefits more from cooperation than the follower.

Our next result shows how the market-size asymmetry affects each country's equilibrium welfare.

Proposition 4. *In all non-cooperative games G^1 , G^2 , G^N and under Nash bargaining, the large country 1's equilibrium welfare is strictly increasing (resp. decreasing) in ε iff $\alpha > \frac{\varepsilon}{3}$ (resp. $\alpha < \frac{\varepsilon}{3}$). The small country 2's equilibrium welfare is strictly decreasing in ε for all parameter values.*

Proof. See Appendix 1.A.5.

Reducing the market size of the small country reduces its tax base and increases the tax base elasticity.²⁴ As a result, the low tax country will tax less and enforce less in equilibrium. On the contrary, increasing the market size of the large country increases its tax base, while

²⁴ According to the definition of ε , an increase in the market-size asymmetry (i.e., a larger value of ε) means an increase (resp. decrease) in country 1's (resp. country 2's) market size.

the welfare effect is ambiguous because of the lower enforcement level from the small country. To understand this, keeping the overall enforcement level $\delta(e_1, e_2)$ unchanged, the marginal rate of substitution is $-\frac{de_1/e_1}{de_2/e_2} = \frac{1-\alpha}{\alpha}$. Hence, as α decreases, country 1 needs to increase its enforcement to a greater extent to offset the enforcement cut of country 2. As Proposition 4 reveals, when α is very small relative to the market asymmetry (i.e., $\alpha < \frac{\varepsilon}{3}$), it becomes too costly for country 1 to offset the enforcement cut of country 2. In this case, increasing the market-size asymmetry also hams the large country.

1.4. Considering tax leadership

We have so far analyzed enforcement leadership and Nash bargaining over enforcements. In this section, we extend our analysis to investigate tax leadership, which can be considered within our framework without many analytical difficulties, and show the robustness of our main results. To this end, we keep the rest of the base model unchanged, and replace stage 3 by another timing game concerning tax decisions: two countries decide whether to move “Early” or “Late” in choosing tax rates; then according to their commitments, two countries non-cooperatively set taxes to maximize their own tax revenues.

The stage 4 equilibrium outcomes are as in Lemma 1(i). Proceeding backward, the tax timing game at stage 3 is solved by computing each country’s equilibrium revenue with different tax leaderships, which is given by:

$$\begin{cases} R_1^{TL} = \frac{(3+\varepsilon)^2}{16} \delta(e_1, e_2) \\ R_2^{TF} = \frac{(5-\varepsilon)^2}{32} \delta(e_1, e_2) \end{cases} \text{ and } \begin{cases} R_1^{TF} = \frac{(5+\varepsilon)^2}{32} \delta(e_1, e_2) \\ R_2^{TL} = \frac{(3-\varepsilon)^2}{16} \delta(e_1, e_2) \end{cases},$$

where superscripts TL and TF respectively denote the tax leader and follower in each sequential game at stage 3.

Thus, we have the following ranking of countries’ tax revenues at stage 3:

$$\forall \varepsilon \in (0, 1], R_i^{TF} > R_i^{TL} > R_i^*, \quad i = 1, 2, \quad (1.8)$$

where $R_i^* = \frac{(3+\mathbf{1}\varepsilon)^2}{18} \delta(e_1, e_2)$ as in Lemma 1(ii).

It directly follows from (1.8) that the tax leadership by either country 1 or country 2 is the stage-3 equilibria and that there is a second-mover advantage for taxes as well as for enforcements. Moreover, it is straightforward to show:

$$\forall \varepsilon \in (0,1], (R_1^{TL} - R_1^*)(R_2^{TF} - R_2^*) > (R_1^{TF} - R_1^*)(R_2^{TL} - R_2^*).$$

This means that the product of deviation losses is larger when the large country is the tax leader. Then according to the definition of risk-dominance, the tax leadership by country 1 (i.e., the large country) is a risk-dominant equilibrium.²⁵ At stage 2, each country's welfare function under country 1's tax leadership is:

$$W_i(e_1, e_2) = \tilde{d}_i e_1^\alpha e_2^{1-\alpha} - \frac{\eta}{2} e_i^2, \quad (1.9)$$

$$\text{where } \tilde{d}_1 = \frac{(3+\varepsilon)^2}{16} \text{ and } \tilde{d}_2 = \frac{(5-\varepsilon)^2}{32} \text{ with } \tilde{d}_1 > d_1.$$

It is easily seen that (1.9) has the same form as (1.3), except for coefficient $\tilde{d}_i$. So all the reasoning in the base model can be applied to solve the enforcement leadership game and Nash bargaining problem when tax leadership is endogenized, which means that Propositions 1, 2 and 3 still hold. On the other hand, according to (A19)²⁶, each country's equilibrium enforcement and welfare levels under Stackelberg games G^1 , G^2 and under bargaining are strictly increasing in d_1 and d_2 . Consequently, compared to simultaneous tax competition in the base model, tax leadership by country 1 strictly increases equilibrium enforcements, taxes and welfare of two countries in the enforcement leadership game and Nash bargaining problem. It is worth noting that the enforcement leadership is driven by the asymmetry in enforcement productivity, whereas the tax leadership is driven by the market-size asymmetry.

1.5. Conclusion

This paper has investigated tax enforcement leadership and cooperation in the context of international profit shifting. The benefit of tax enforcement is to raise the cost of profit shifting for multinationals, which makes tax bases less sensitive to tax rates enabling countries to sustain higher corporate taxes. We depart from Hindriks and Nishimura (2021) in which enforcement decisions are simultaneous, to allow for the possible leadership in tax

²⁵ Tax leadership by the large country is in line with the existing literature on tax competition (see, e.g., Hindriks and Nishimura, 2015, 2017). All the detailed derivations involved in stage-3 equilibrium are available upon request.

²⁶ See Appendix 1.A.5.

enforcement. Enforcement leadership is consistent with the evidence that progresses of strengthening tax enforcements are different among countries.

To endogenize the timing of enforcement decisions, the paper considers an “enforcement leadership game” in which countries commit themselves to move early or late in choosing their enforcement levels. Our model incorporates two types of asymmetry between countries. Besides the asymmetry in market size, countries are also asymmetric in enforcement productivity (i.e., enforcement elasticities of tax revenues). The Subgame Perfect Equilibria (SPEs) of the enforcement leadership game consists of two Stackelberg equilibria at which two countries set enforcements sequentially. Both countries are better off by setting their enforcement sequentially instead of simultaneously. We then resort to the risk dominance criterion to solve for the coordination issue caused by multiple equilibria, and establish that enforcement leadership by the country with low enforcement productivity risk dominates. Therefore, our model finds a novel driving force of enforcement leadership (the asymmetric enforcement productivity), which is distinct from that of tax leadership (the asymmetric market size).

A key feature of our model is that enforcement is chosen before taxes. Enforcements act as a level playing field, before countries compete in taxes. The implication is that even the low-tax country can benefit from higher enforcement. The positive enforcement externality leads to suboptimal levels of non-cooperative enforcement. To reach Pareto efficiency, we consider enforcement cooperation as a Nash Bargaining problem. We establish that both countries’ enforcement levels, taxes and welfare levels are strictly higher under Nash bargaining than under enforcement leadership. The country with a low-enforcement productivity (also the enforcement leader) benefits more from the cooperation (both in terms of welfare gain and welfare share).

We then show the robustness of our results to endogenous tax leadership. We confirm in our framework that the tax leadership by the large country risk dominates and that it can increase two countries’ equilibrium enforcement, tax and welfare levels under both cooperative and non-cooperative enforcement.

A key consideration in the international taxation challenge is the enforcement decisions and cooperation. Our results suggest that the enforcement leadership and the benefit of enforcement cooperation closely relate to the asymmetry in enforcement productivity. This indicates the importance of an empirical analysis of countries’ enforcement elasticities of tax revenue in the context of international profit shifting.

Appendix 1.A. Proofs.

1.A.1. Proof of Proposition 1:

Given the results in Table 1.1, we can compute the product of deviation losses:

$$\Pi = \frac{\alpha^{2\alpha}(1-\alpha)^{2-2\alpha}}{4\eta^2(2-\alpha)^2} d_1^{1+2\alpha} d_2^{3-2\alpha} J(\alpha),$$

where $J(\alpha) := 2^{-\alpha}[2^{1-\alpha}(2-\alpha)^\alpha + \alpha - 2][2^\alpha(2-\alpha)^2(1+\alpha) - 4(2-\alpha)^\alpha] \\ + (2-\alpha)^2(1+\alpha)^{-2\alpha}[2^\alpha - (1+\alpha)^\alpha][2^{1+\alpha} - (2-\alpha)(1+\alpha)^{1+\alpha}].$

The graph of function $J(\alpha)$ over interval $[0,1]$ is:²⁷

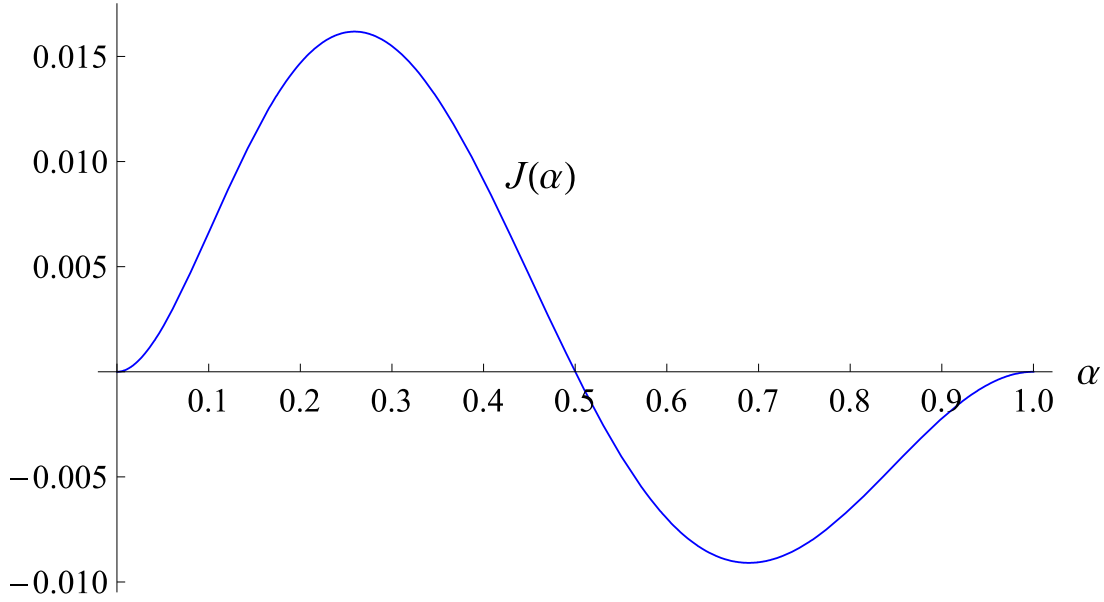


Fig. A1. The graph of $J(\alpha)$

From Fig. A1, we have:

$$\begin{cases} \Pi > 0 & \text{if } 0 < \alpha < \frac{1}{2}, \\ \Pi < 0 & \text{if } \frac{1}{2} < \alpha < 1 \end{cases},$$

which immediately leads to Proposition 1(i).

²⁷ Figs. A1, A2 and A3 are plotted by Wolfram Mathematica. The corresponding programs are available upon request.

The second part of the proposition directly follows from (1.4). **Q.E.D.**

1.A.2. Proof of Lemma 2:

We analyze game G^i in which country i leads in the enforcement competition and country j follows. As for country i , note that:

$$\left. \frac{\partial W_i(k e_i^L, k e_j^F)}{\partial k} \right|_{k=1} = \frac{\partial W_i(e_i^L, e_j^F)}{\partial e_i^L} e_i^L + \frac{\partial W_i(e_i^L, e_j^F)}{\partial e_j^F} e_j^F. \quad (\text{A1})$$

The first-order condition of country i 's welfare optimization problem reads:

$$\frac{\partial W_i(e_i^L, e_j^F(e_i^L))}{\partial e_i^L} = \frac{\partial W_i(e_i^L, e_j^F)}{\partial e_i^L} + \frac{\partial W_i(e_i^L, e_j^F)}{\partial e_j^F} \frac{de_j^F(e_i^L)}{de_i^L} = 0. \quad (\text{A2})$$

Given (A1), (A2), (1.3) and (1.5), we have: $\forall \alpha \in (0,1)$,

$$\left. \frac{\partial W_i(k e_i^L, k e_j^F)}{\partial k} \right|_{k=1} = \frac{\partial W_i(e_i^L, e_j^F)}{\partial e_j^F} e_j^F \left(1 - \frac{de_j^F / e_j^F}{de_i^L / e_i^L}\right) = d_i \frac{\partial \delta(e_i^L, e_j^F)}{\partial e_j^F} e_j^F (1 - \varepsilon_j) > 0,$$

where $\varepsilon_j := \frac{de_j^F / e_j^F}{de_i^L / e_i^L} = \begin{cases} (1-\alpha)/(2-\alpha) & \text{if } j=1, \\ \alpha/(1+\alpha) & \text{if } j=2. \end{cases}$ is leader i 's enforcement elasticity of follower j 's enforcement.

Thus, for any $\alpha \in (0,1)$ there exists some $\varepsilon_1 > 0$ such that $\forall k \in (1, 1+\varepsilon_1)$,

$$W_i(k e_i^L, k e_j^F) > W_i(e_i^L, e_j^F) := W_i^L.$$

For country j , we have:

$$\left. \frac{\partial W_j(k e_i^L, k e_j^F)}{\partial k} \right|_{k=1} = \frac{\partial W_j(e_i^L, e_j^F)}{\partial e_i^L} e_i^L + \frac{\partial W_j(e_i^L, e_j^F)}{\partial e_j^F} e_j^F = \frac{\partial W_j(e_i^L, e_j^F)}{\partial e_i^L} e_i^L = d_j \frac{\partial \delta(e_i^L, e_j^F)}{\partial e_i^L} e_i^L > 0,$$

where the second equality derives from the first-order condition of country j 's welfare optimization problem and the third equality is obtained from (1.3).

This implies that for any $\alpha \in (0,1)$ there exists some $\varepsilon_2 > 0$ such that $\forall k \in (1, 1+\varepsilon_2)$,

$$W_j(ke_i^L, ke_j^F) > W_j(e_i^L, e_j^F) := W_j^F.$$

Take $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$, then $\forall k \in (1, 1+\varepsilon)$, $\begin{cases} W_i(ke_i^L, ke_j^F) > W_i^L \\ W_j(ke_i^L, ke_j^F) > W_j^F \end{cases}$ ($i, j = 1, 2; i \neq j$), i.e.,

increasing two countries' enforcement levels by $k-1$ percent from the Stackelberg levels can make both countries better off. **Q.E.D.**

1.A.3. Proof of Proposition 2:

We analyze Case (i) $\alpha \in (0, \frac{1}{2})$ and Case (ii) $\alpha \in (\frac{1}{2}, 1)$, separately.

Case (i): $\alpha \in (0, \frac{1}{2})$. In this case, the disagreement outcome corresponds to Stackelberg game G^1 .

Lemma 2 implies that each country's welfare under Nash bargaining is strictly higher than in Stackelberg game G^1 , i.e., $\begin{cases} W_1(e_1^B, e_2^B) > W_1(e_1^L, e_2^F) \\ W_2(e_1^B, e_2^B) > W_2(e_1^L, e_2^F) \end{cases}$.

Now we turn to the enforcement levels. Firstly, we show that $e_1^B > e_1^L$. The proof proceeds by contradiction.

If $e_1^L \geq e_1^B$, then $W_2(e_2^B, e_1^L) \geq W_2(e_2^B, e_1^B)$ since $W_2(e_2, e_1)$ is increasing in e_1 .

On the other hand, according to the definition of the Stackelberg equilibrium of game G^1 , we have:

$$W_2(e_2^F, e_1^L) \equiv \max_{e_2} W_2(e_2, e_1^L) \geq W_2(e_2^B, e_1^L).$$

As a result, $W_2(e_2^F, e_1^L) \geq W_2(e_2^B, e_1^B)$, which violates the fact $W_2(e_2^B, e_1^B) > W_2(e_2^F, e_1^L)$.

Therefore, $e_1^B > e_1^L$ always holds.

Now we prove that $e_2^B > e_2^F$, also proceeding by contradiction.

If $e_2^F \geq e_2^B$, then the mean value theorem gives:

$$W_1(e_1^B, e_2^B) - W_1(e_1^L, e_2^F) = \frac{\partial W_1(e_1', e_2')}{\partial e_1'}(e_1^B - e_1^L) + \frac{\partial W_1(e_1', e_2')}{\partial e_2'}(e_2^B - e_2^F), \quad (\text{A3})$$

where $e_1^L < e_1' < e_1^B$ and $e_2^B \leq e_2' \leq e_2^F$.

On the other hand, notice that:

$$\frac{\partial W_1(e_1', e_2')}{\partial e_1'} \leq \frac{\partial W_1(e_1', e_2^F)}{\partial e_1'} < \frac{\partial W_1(e_1^L, e_2^F)}{\partial e_1^L} < 0, \quad (\text{A4})$$

where the first inequality holds since $W_1(e_1, e_2)$ is super-modular, the second holds since $W_1(e_1, e_2)$ is concave w.r.t. e_1 ,²⁸ and the last follows from the first-order condition of the welfare maximization problem for country 1 in game G^1 .²⁹

Combining (A3) and (A4) yields:

$$W_1(e_1^B, e_2^B) - W_1(e_1^L, e_2^F) \leq \frac{\partial W_1(e_1', e_2')}{\partial e_1'}(e_1^B - e_1^L) < 0, \quad (\text{A5})$$

where the first inequality holds because of the positive externality of enforcement $\frac{\partial W_1(e_1', e_2')}{\partial e_2'} > 0$.

(A5) contradicts country 1's welfare ranking $W_1(e_1^B, e_2^B) > W_1(e_1^L, e_2^F)$.

Thus, it must be that $e_2^B > e_2^F$.

Then the comparison of tax rates immediately follows from the comparison of enforcement levels and Lemma 1(ii).

Case (ii): $\alpha \in (\frac{1}{2}, 1)$. The disagreement outcome corresponds to Stackelberg game G^2 .

By the similar reasoning used in Case (i), it is straightforward to show that Proposition 2 still holds. **Q.E.D.**

²⁸ Given (1.2), the super-modularity and concavity of the enforcement technology (here, it is the Cobb-Douglas function) respectively lead to the super-modularity and concavity of two countries' welfare functions.

²⁹ From (A2), we have: $\frac{\partial W_1(e_1^L, e_2^F)}{\partial e_1^L} = - \underbrace{\frac{\partial W_1(e_1^L, e_2^F)}{\partial e_2^F}}_{+} \cdot \underbrace{\frac{de_2^F(e_1^L)}{de_1^L}}_{+} < 0$.

1.A.4. Proof of Proposition 3:

Proposition 2 implies that all the constraints in Problem (1.7) are non-binding at the bargaining solution. Therefore, the first-order conditions of Problem (1.7) are given by:

$$\begin{cases} \frac{\partial W_1(e_1, e_2)}{\partial e_1} [W_2(e_1, e_2) - W_2^v] + \frac{\partial W_2(e_1, e_2)}{\partial e_1} [W_1(e_1, e_2) - W_1^u] = 0 \\ \frac{\partial W_1(e_1, e_2)}{\partial e_2} [W_2(e_1, e_2) - W_2^v] + \frac{\partial W_2(e_1, e_2)}{\partial e_2} [W_1(e_1, e_2) - W_1^u] = 0 \end{cases}. \quad (\text{A6})$$

From (A6), we obtain:

$$\frac{\partial W_1(e_1, e_2)}{\partial e_1} \frac{\partial W_2(e_1, e_2)}{\partial e_2} = \frac{\partial W_1(e_1, e_2)}{\partial e_2} \frac{\partial W_2(e_1, e_2)}{\partial e_1}.$$

Given (1.3), the above equation can be simplified as:

$$e_1 e_2 = \frac{1}{\eta} [\alpha d_1 e_2 \left(\frac{e_1}{e_2}\right)^{\alpha-1} + (1-\alpha) d_2 e_1 \left(\frac{e_1}{e_2}\right)^{\alpha}]. \quad (\text{A7})$$

Recall that the disagreement point relates to the value of α : (i) $0 < \alpha < \frac{1}{2}$ or (ii) $\frac{1}{2} < \alpha < 1$. We now analyze each case separately.

Case (i): $0 < \alpha < \frac{1}{2}$. According to Proposition 1(i), country 1 leads if the negotiation collapses. In line with this, we define:

$$\begin{cases} e_1 := k e_1^L \\ e_2 := k h e_2^F \end{cases}, \quad (\text{A8})$$

where variables $k > 0$ and $h > 0$.

Using (A8), we can transform (A6) and (A7) into equations w.r.t. k and h . Specifically, plugging (A8) and the values of e_1^L , e_2^F shown in Table 1.1 into (A7), with tedious calculations, yields the relationship of k and h :

$$k = \frac{1}{2} h^{-1-\alpha} [(1+\alpha)h^2 + 2]. \quad (\text{A9})$$

Substituting (A8) and (A9) into (A6), through tedious computation and simplification,

we can obtain that for any $\alpha \in (0, \frac{1}{2})$, h is the solution to the following equation:³⁰

$$4 - 4h^2 - 3h^4 + 2\alpha^3 h^4 + 4h^{2\alpha}(h^2 - 1) + \alpha^2 h^2(h^2 + 8) - 4\alpha(h^4 - h^2 + h^{2\alpha} - 2) = 0. \quad (\text{A10})$$

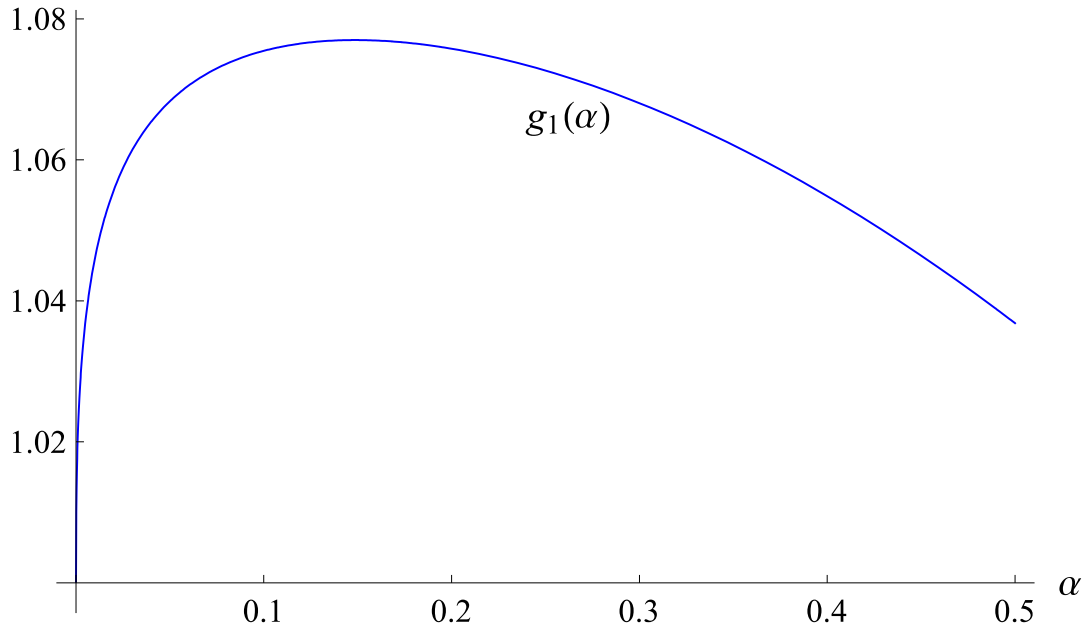
Plugging (A9) and the values of e_1^L and e_2^F into (1.2), we get two countries' welfare W_i^B ($i = 1, 2$) under Nash bargaining:

$$\begin{cases} W_1^B = \frac{1}{4} h^{-2(1+\alpha)} [(2-\alpha)(1+\alpha)^2 h^4 + 4(1-\alpha^2)h^2 - 4\alpha] W_1^L \\ W_2^B = \frac{1}{4} h^{-2\alpha} [4 + 4\alpha h^2 + (\alpha^2 - 1)h^4] W_2^F \end{cases}, \quad (\text{A11})$$

which yields:

$$\frac{W_1^B / W_1^L}{W_2^B / W_2^F} = -\frac{\alpha}{h^2} + \frac{2}{2 + (\alpha - 1)h^2} =: g_1(\alpha). \quad (\text{A12})$$

Given (A10) and (A12), we can graph function $g_1(\alpha)$ over interval $[0, \frac{1}{2}]$ as below:



³⁰ For any $\alpha \in (0, \frac{1}{2})$ (resp. $\alpha \in (\frac{1}{2}, 1)$), we can show that there exists a unique h with $h > 0$ satisfying (A10) (resp. (A15)). Therefore, h is an implicit function of α . To save space, we omit the tedious proof of this result, which is available upon request.

Fig. A2. The graph of $g_1(\alpha)$

According to Fig. A2, we have:

$$\forall \alpha \in (0, \frac{1}{2}), \quad \frac{W_1^B / W_1^L}{W_2^B / W_2^F} > 1. \quad (\text{A13})$$

Case (ii): $\frac{1}{2} < \alpha < 1$. Country 2 leads if the negotiation collapses. Then we define:

$$\begin{cases} e_1 := ke_1^F \\ e_2 := khe_2^L \end{cases}, \quad (\text{A14})$$

where variables $k, h > 0$.

Following the same steps as in Case (i), we can get the relationship between k and h in Case (ii): $k = \frac{1}{2}h^{-1-\alpha}(2-\alpha+2h^2)$.

For any $\alpha \in (\frac{1}{2}, 1)$, h is the solution to the following equation:

$$4 + 4\alpha - 7\alpha^2 + 2\alpha^3 - (8 - 20\alpha + 8\alpha^2)h^2 - (12 - 8\alpha)h^4 - 4h^{2\alpha}(1 + (\alpha - 2)h^2) = 0. \quad (\text{A15})$$

Two countries' cooperative welfare levels W_i^B ($i = 1, 2$) are:

$$\begin{cases} W_1^B = \frac{1}{4}h^{-2(1+\alpha)}(\alpha - 2h^2)[\alpha - 2(1 + h^2)]W_1^F \\ W_2^B = \frac{1}{4}h^{-2\alpha}[4 + (\alpha - 2h^2)((\alpha - 3)\alpha - 2(\alpha - 1)h^2)]W_2^L \end{cases}, \quad (\text{A16})$$

which leads to:

$$\frac{W_1^B / W_1^F}{W_2^B / W_2^L} = \frac{2h^2 - \alpha}{h^2[2 + \alpha - \alpha^2 + 2(\alpha - 1)h^2]} =: g_2(\alpha). \quad (\text{A17})$$

Likewise, given (A15) and (A17), we can graph function $g_2(\alpha)$ over interval $[\frac{1}{2}, 1]$:

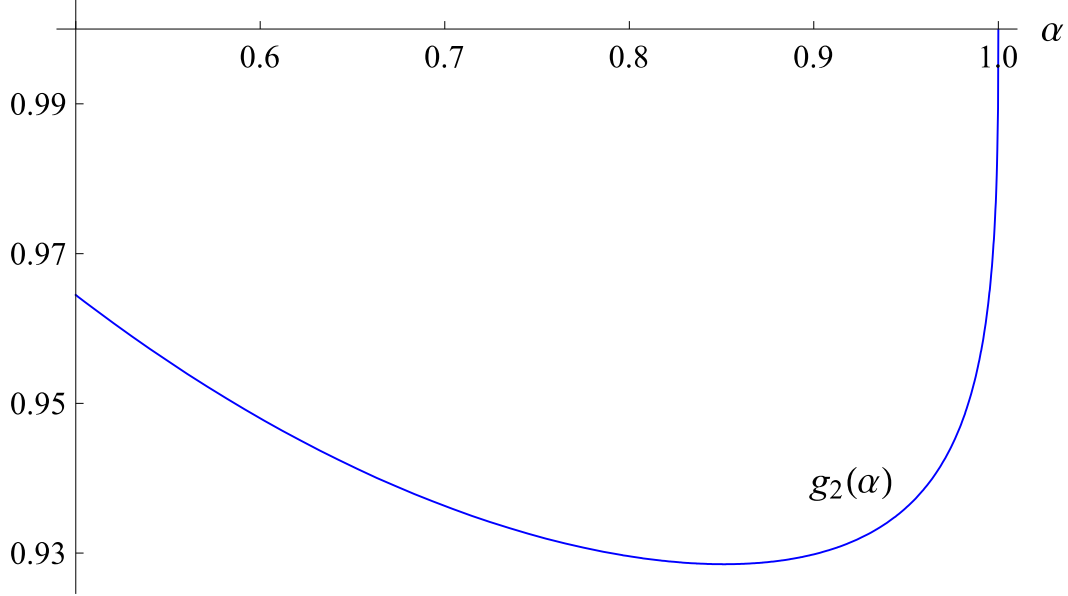


Fig. A3. The graph of $g_2(\alpha)$

From Fig. A3, we get:

$$\forall \alpha \in \left(\frac{1}{2}, 1\right), \quad \frac{W_1^B / W_1^F}{W_2^B / W_2^L} < 1. \quad (\text{A18})$$

Combining (A13) and (A18) yields Proposition 3. **Q.E.D.**

1.A.5. Proof of Proposition 4:

Recall that h , independent of ε , is a function of α . Then given (A8), (A11), (A14), (A16) and the results shown in Table 1.1, two countries' equilibrium enforcement and welfare levels in games G^1 , G^2 , G^N and in bargaining problem display the following form:

$$\begin{cases} e_1^u = F_1^u(\alpha) d_1^{(1+\alpha)/2} d_2^{(1-\alpha)/2} \\ e_2^u = F_2^u(\alpha) d_1^{\alpha/2} d_2^{(2-\alpha)/2} \end{cases} \quad \text{and} \quad \begin{cases} W_1^u = H_1^u(\alpha) d_1^{1+\alpha} d_2^{1-\alpha} \\ W_2^u = H_2^u(\alpha) d_1^\alpha d_2^{2-\alpha} \end{cases}, \quad (\text{A19})$$

where $u = N, L, F, B$, and $F_i^u(\alpha)$, $H_i^u(\alpha)$ ($i = 1, 2$) are functions of α and are independent of parameter ε .

Given (A19), the comparative static effects of ε under three non-cooperative games and under bargaining are given by:

$$\begin{cases} \frac{\partial W_1^u}{\partial \varepsilon} = \left(\frac{1+\alpha}{d_1} \frac{dd_1}{d\varepsilon} + \frac{1-\alpha}{d_2} \frac{dd_2}{d\varepsilon} \right) W_1^u = \frac{4W_1^u}{9-\varepsilon^2} (3\alpha - \varepsilon) \\ \frac{\partial W_2^u}{\partial \varepsilon} = \left(\frac{\alpha}{d_1} \frac{dd_1}{d\varepsilon} + \frac{2-\alpha}{d_2} \frac{dd_2}{d\varepsilon} \right) W_2^u = \frac{4W_2^u}{9-\varepsilon^2} (3\alpha - \varepsilon - 3) \end{cases}, \text{ where } u = N, L, F, B.$$

Then the proposition immediately follows. **Q.E.D.**

Chapter 2. Multinational Taxation Under Pressure: The Role of Tax Deductibility¹

Abstract

To address 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 MNEs' behavior and countries' corporate taxes in a formal model of corporate tax competition with two countries and a tax haven. A profit tax with full deductibility (under either separate accounting (SA) or formula apportionment (FA)) preserves production efficiency but incurs tax base erosion. A turnover tax with no deductibility eliminates profit shifting incentives but causes production distortion. Besides, we identify a formula manipulation effect for profit tax under FA, which negatively affects countries' taxes. We show that profit tax under FA dominates when the output elasticity is high and the tax capacity is intermediate. Otherwise, turnover tax dominates for low tax capacity, and profit tax under SA dominates for high tax capacity. We then analyze the general case of tax deductibility and revisit the Diamond-Mirrlees production efficiency theorem in the profit shifting context. We show that a lower tax deductibility can increase tax revenue under sufficiently low tax capacity. By contrast, equilibrium taxes strictly increase with tax deductibility when tax capacity is high, marginal production cost is low or the output elasticity is high.

JEL classification: H25, H71, H73, F23, D24

Keywords: Corporate taxes, Profit shifting, Tax haven, Tax deductibility, Tax competition, Digital economy, Extractive sector

¹ This chapter is co-authored with Prof. Jean Hindriks.

2.1. Introduction

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., 2023). Base erosion and profit shifting (BEPS) persists and raises problems for corporate 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. In particular, the European Commission (2023) launched the “Business in Europe: Framework for Income Taxation (BEFIT)” proposal which aims to replace SA with FA as the method of multinational taxation in the EU.² 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; Riedel and Runkel, 2007; Nielsen et al., 2010; Gresik, 2016; Mardan and Stimmelmayer, 2018). In this regard, our work is most closely related to Riedel and Runkel (2007), who investigate a tax

² The BEFIT proposal is a renewal of the Commission's “Common Consolidated Corporate Tax Base (CCCTB)” proposal in 2016. The CCCTB proposal was withdrawn, since it failed to gain sufficient political support among EU Member States.

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.³ 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. By contrast, under high tax capacity, low marginal production cost or large output elasticity, the benefit from increasing production efficiency is first-order, and countries' equilibrium taxes are strictly increasing in tax deductibility. In

³ As in most theoretical literature on profit shifting, we assume that the MNE is a price taker in the goods market (see Section 2.2), which implies that the demand for the goods is perfectly elastic. Therefore, the MNE is unable to pass the tax on to consumers and so the tax incidence falls entirely on the MNE. In this setting, the consequence of deviating from pure profit tax is the output distortion, and the tax effect on the consumption side is absent. See Best et al. (2015) for two additional effects (cascading and incidence effects) in the general equilibrium analysis of firm taxation.

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 2.2 sets up a model of international tax competition with two countries and a tax haven. Section 2.3 characterizes equilibrium outcomes in different tax regimes – a pure profit tax with either SA or FA and a turnover tax. Section 2.4 compares the three tax regimes in terms of equilibrium tax revenues. In Section 2.5, we go beyond the polar cases – profit and turnover taxation – to investigate the general case of tax deductibility. Section 2.6 concludes with a few policy remarks.

2.2. The model

Consider an economy with two identical countries labelled by 1, 2 and a tax haven indexed by 3. Each of them hosts an affiliate of a representative multinational enterprise (MNE).⁴ 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 $\alpha > 0$, $0 \leq \beta < 1$ 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

⁴ In our model, the three affiliates are set up by the MNE ex ante, and their locations are assumed to be fixed. In contrast, the profits are mobile among the three affiliates. The exogeneity of firms' location is a usual assumption in the theoretical works on profit shifting. This may reflect the fact that it is much easier and less costly to reallocate paper profits than to relocate firms. In some profit shifting literature (e.g., Peralta et al., 2006; Bucovetsky and Haufler, 2008), the mobility of MNEs is considered in tax competition games and MNEs decide their location in response to countries' corporate tax rates.

focus on technological parameters α and k in the convex component of $c(y_i)$. In Appendix 2.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 haven affiliate has no real production.

The MNE can shift profits from production affiliates to the tax haven via transfer pricing process (see Dischinger and Riedel, 2011; Faulkender et al., 2019). To do so, an intangible asset is located in the tax haven affiliate, which charges transfer prices to affiliates in two countries for using the intangible input. In addition, the MNE may also shift profits between non-haven countries (see Johannesen, 2010; Álvarez-Martínez et al., 2022). We assume without loss of generality that the affiliate in country 2 owns a second intangible asset which is essential for production. Affiliate 1 accesses this intangible asset by paying a transfer price to affiliate 2. In this set-up, the MNE can reallocate its profits among the three affiliates so as to minimize its tax liability. Denote by g_{ij} the intra-firm transfer price paid by affiliate i to j ($1 \leq i < j \leq 3$). For simplicity, the arm's length prices for the intangible assets are normalized to zero. Then $g_{ij} > 0$ ($g_{ij} < 0$) represents that the transfer price is overstated (understated), and hence profit is shifted to (from) affiliate j .

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 δ captures countries' tax capacity (see, e.g., Kind et al., 2005; Devereux et al., 2008; Mardan and Stimmelmayer, 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 (zero-tax location), 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 tax 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:

$$\begin{aligned} \pi_i &= y_i - c(y_i) - \mathbf{1}_i g_{12} - g_{i3} - t_i \pi_i^t, \quad i = 1, 2, \\ \pi_3 &= \sum_{i=1}^2 g_{i3}, \end{aligned} \tag{2.1}$$

where the indicator variable $\mathbf{1}_i$ equals unity for $i=1$ and -1 for $i=2$, and 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 transfer prices to maximize total after-tax profits.

2.3. 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.

2.3.1. 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:

$$\pi_i^t = y_i - c(y_i) - \mathbf{1}_i g_{12} - g_{i3}, \quad i = 1, 2. \quad (2.2)$$

Given (2.1) and (2.2), summing up the three affiliates' net-of-tax profits minus the concealment costs lead to the total after-tax profit of the MNE:

$$\Pi^{SA} = \sum_{i=1}^2 \left[(1-t_i)(y_i - c(y_i) - \mathbf{1}_i g_{12} - g_{i3}) \right] + \sum_{i=1}^2 g_{i3} - \sum_{1 \leq i < j \leq 3} \frac{\delta}{2} g_{ij}^2. \quad (2.3)$$

The MNE chooses output y_i and transfer price g_{ij} ($1 \leq i < j \leq 3$) to maximize (2.3). Solving the first-order conditions of the MNE's profit maximization gives:

$$c'(y_i) = 1, \quad i = 1, 2, \quad (2.4)$$

$$g_{12} = \frac{t_1 - t_2}{\delta}, \quad g_{i3} = \frac{t_i}{\delta}, \quad i = 1, 2. \quad (2.5)$$

From (2.4), 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 $\alpha \leq \frac{1-\beta}{e}$. In this paper, we keep the assumption

$0 < \alpha \leq \frac{1-\beta}{e}$, which is used for Proposition 1(ii), Proposition 2(iii) and Proposition 3(ii)

with respect to parameter k .

Given (2.2), (2.4) and (2.5), country i 's tax base loss from a marginal increase in its tax

rate is $\frac{\partial(\mathbf{1}_i g_{12} + g_{i3})}{\partial t_i} = \frac{2}{\delta}$. 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 (2.2). In doing so, it takes into account the MNE's profit maximizing solutions which depend on t_i and t_j according to (2.4) and (2.5). Solving the first-order conditions of tax revenue maximization, we can get the Nash equilibrium tax rate and revenue of each country:

$$\tilde{t}(\delta) = \min \left\{ \frac{\bar{\pi}\delta}{3}, 1 \right\}, \quad (2.6)$$

$$\tilde{R}(\delta) = \begin{cases} \frac{2\bar{\pi}^2\delta}{9} & \text{if } 0 < \delta < \frac{3}{\bar{\pi}} \\ \bar{\pi} - \frac{1}{\delta} & \text{if } \delta \geq \frac{3}{\bar{\pi}} \end{cases}. \quad (2.7)$$

Given (2.6) and (2.7), 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 $\delta \geq \frac{3}{\pi}$; (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.

2.3.2. 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 $\lambda_i := \frac{y_i}{\sum_{j=1}^2 y_j}$. Using (2.2), the consolidated tax base in the FA

union is:

$$\pi^c := \sum_{i=1}^2 [y_i - c(y_i) - \mathbf{1}_i g_{i2} - g_{i3}] = \sum_{i=1}^2 [y_i - c(y_i) - g_{i3}]. \quad (2.8)$$

The MNE's taxable income in country i reads:

$$\pi_i^t = \lambda_i \pi^c, \quad i = 1, 2. \quad (2.9)$$

Given (2.1), (2.8) and (2.9), the MNE's total after-tax profit under FA reads:

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

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:

$$\frac{\partial \Pi^{FA}}{\partial g_{12}} = -\delta g_{12} = 0, \quad (2.10)$$

$$\frac{\partial \Pi^{FA}}{\partial g_{i3}} = \tau - \delta g_{i3} = 0, \quad i = 1, 2, \quad (2.11)$$

$$\frac{\partial \Pi^{FA}}{\partial y_i} = -\frac{\partial \tau}{\partial y_i} \pi^c + (1-\tau) \frac{\partial \pi^c}{\partial y_i} = \frac{(\tau - t_i) \pi^c}{\sum_{i=1}^2 y_i} + (1-\tau)[1 - c'(y_i)] = 0, \quad i = 1, 2. \quad (2.12)$$

Because of the tax base consolidation within the FA union, (2.10) 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 (2.11), 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 (2.12) 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 (2.11) and (2.12), together with the symmetry property, yields the comparative statics under identical tax rate $t_i = t$ ($i = 1, 2$):

$$\frac{\partial g_{i3}}{\partial t_i} = \frac{\partial g_{j3}}{\partial t_i} = \frac{1}{2\delta} > 0, \quad i, j = 1, 2, \quad (2.13)$$

$$\frac{\partial y_i}{\partial t_i} = -\frac{\bar{\pi} - t / \delta}{2(1-t)\bar{y}c''(\bar{y})} < 0, \quad i = 1, 2, \quad (2.14)$$

$$\frac{\partial y_j}{\partial t_i} = \frac{\bar{\pi} - t / \delta}{2(1-t)\bar{y}c''(\bar{y})} > 0, \quad i, j = 1, 2; \quad i \neq j. \quad (2.15)$$

Under FA, the MNE's response to one country's tax increase is twofold. Firstly, as shown by (2.13), 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 (2.14) and (2.15), 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 (2.8). In doing so, it takes into account the MNE's production and profit shifting choices which rely on countries' tax rates according to (2.10) – (2.12). 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:

$$\frac{\partial R_i}{\partial t_i} = \bar{\pi} - \frac{3\hat{t}}{2\delta} - \frac{\hat{t}(\bar{\pi} - \frac{\hat{t}}{\delta})^2}{2(1-\hat{t})\bar{y}^2 c''(\bar{y})} = 0. \quad (2.16)$$

(2.16) determines two countries' equilibrium tax rates. Due to the asymmetry between FA countries and the tax haven, the solution to (2.16) 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 2.A.1.

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 (2.13) – (2.15) and the symmetry condition, we compute the loss of one country's tax base from increasing its own tax rate:

$-\frac{\partial \pi_i^t}{\partial t_i} \Big|_{t_i=t_j=t} = \lambda_i \cdot \frac{\partial}{\partial t_i} \sum_{j=1}^2 g_{j3} - \frac{\partial \lambda_i}{\partial t_i} \cdot \pi^c$. 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.

2.3.3. Turnover taxation

Consider now two countries apply a turnover tax regime.⁵ Then the MNE’s taxable income in country i is the gross revenue (i.e., turnover) of affiliate i :

$$\pi_i^t = y_i, \quad i = 1, 2. \quad (2.17)$$

Given (2.1) and (2.17), the MNE’s total after-tax profit under turnover tax amounts to:

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

⁵ Different from the pure profit tax, the two alternative principles (SA and FA) under turnover tax lead to the same outcome. This can be easily verified by letting $\mu = 0$ in Section 2.5 and in Appendix 2.B.

The MNE chooses output y_i and transfer price g_{ij} ($1 \leq i < j \leq 3$) 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:

$$g_{ij} = 0, \quad 1 \leq i < j \leq 3. \quad (2.18)$$

As shown by (2.18), 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:

$$\frac{\partial \Pi^T}{\partial y_i} = 1 - t_i - c'(y_i) = 0, \quad i = 1, 2. \quad (2.19)$$

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

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 (2.19). Solving the first-order condition of the maximization problem yields the equilibrium turnover tax rate:

$$t^T = \frac{(1 - \beta)(k - 1)}{k}. \quad (2.20)$$

Given (2.19) and the specification of $c(y_i)$, following Best et al. (2015), the output elasticity is: $\varepsilon_i := \frac{dy_i / y_i}{d(1 - t_i) / (1 - t_i)} = \frac{1 - t_i}{(k - 1)(1 - t_i - \beta)}$, 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 $k \rightarrow 1$ to $1 - \beta$ as $k \rightarrow +\infty$.

Using (2.19) and (2.20) yields the equilibrium output and revenue of each country:

$$\underline{y} = k^{\frac{1}{1-k}} \bar{y}, \quad (2.21)$$

$$R^T = k^{\frac{1}{1-k}} \bar{\pi}. \quad (2.22)$$

In (2.21) and (2.22), 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 (2.21), the output gap between profit and turnover tax is $\bar{y} - \underline{y} = \left(1 - k^{\frac{1}{1-k}}\right) \bar{y}$, 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 present the main result in the lemma below.

Lemma 3. *Under turnover tax, (i) profit shifting is completely eliminated and both 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.*

Before comparing the equilibrium taxes associated with the three tax regimes, we summarize by the following table the real production and profit shifting responses of the MNE.

Table 2.1: The MNE's behavior under three tax regimes.

Tax Regime		Production	Profit Shifting	Other Effect
Profit tax	SA	efficient	among three affiliates	n/a
	FA	efficient	partially eliminated	formula manipulation effect
Turnover tax		inefficient	eliminated	n/a

As shown by Table 2.1, for pure profit tax with SA, the MNE's output choice is efficient, while it can shift profits among the three affiliates. By contrast, turnover tax causes production inefficiency but eliminates profit shifting. The intermediate case is pure profit tax with FA. It maintains efficient production, removes profit shifting incentive between FA countries, but induces a formula manipulation effect.

2.4. Comparison of the tax regimes

2.4.1. Profit versus turnover taxation

As we have shown, a tax regime switch 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 $\delta < \delta^{FA}$ ($\delta > \delta^{FA}$). FA is always superior to turnover taxation in the absence of the tax haven. (ii) Both tax capacity thresholds δ^{SA} and δ^{FA} strictly increase with technological parameters α and k .*

Proof. See Appendix 2.A.2.

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.

2.4.2. 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 2.A.3.

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 2.A.4.

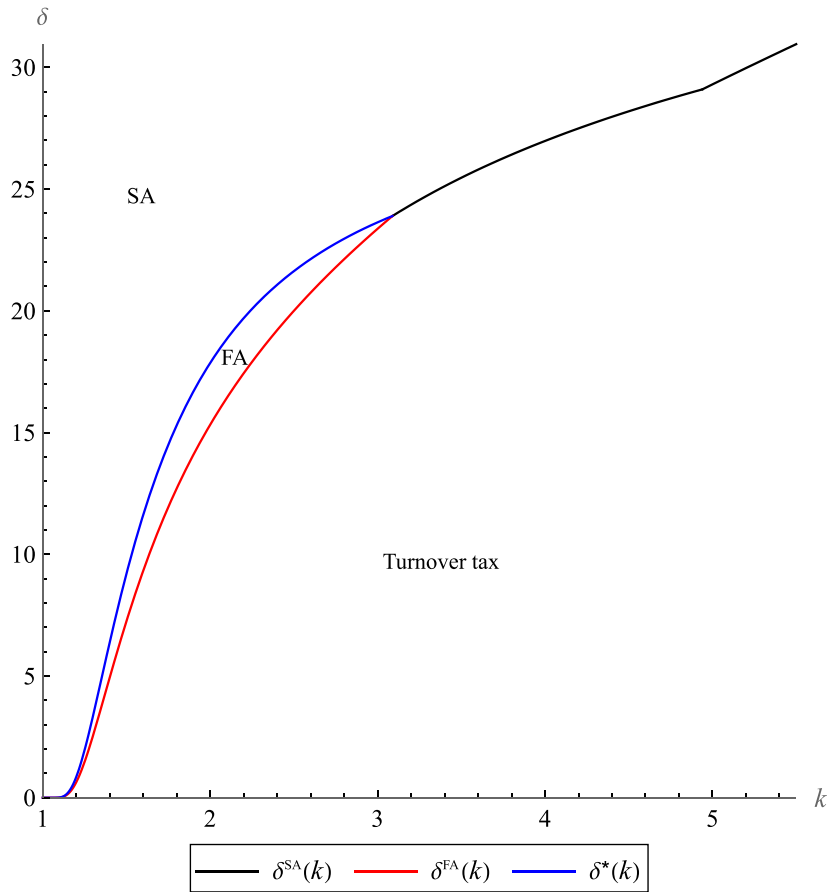


Fig. 2.1. The optimal choice among three tax regimes. Parameter values are specified as: $\alpha = \frac{2}{55}$ and $\beta = \frac{9}{10}$.

Fig. 2.1 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 $k \rightarrow 1$. 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:

$$\frac{\bar{y} - c(\bar{y})}{\bar{y}} = (1 - \beta)(1 - \frac{1}{k}), \text{ which means that the pre-tax profit margin strictly increases with}$$

k . In the mining sector, both marginal production costs and profit margins are relatively high⁶, 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

⁶ See https://pages.stern.nyu.edu/~adamodar/New_Home_Page/data.html.

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 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.

2.5. 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 variable 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 2.B.

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

$$\pi_i^t = y_i - \mu(c(y_i) + \mathbf{1}_i g_{12} + g_{i3}). \quad (2.23)$$

Given (2.1) and (2.23), the MNE's total net-of-tax profit amounts to:

$$\Pi^{SA} = \sum_{i=1}^2 (1-t_i) [y_i - \mu(c(y_i) + \mathbf{1}_i g_{12} + g_{i3})] - (1-\mu) \sum_{i=1}^2 c(y_i) + \mu \sum_{i=1}^2 g_{i3} - \sum_{1 \leq i < j \leq 3} \frac{\delta}{2} g_{ij}^2. \quad (2.24)$$

In the second stage, the MNE chooses output levels and intra-firm transfer prices to maximize (2.24). Solving the first-order conditions of the profit maximization yields:

$$c'(y_i) = 1 - m_i, \quad i = 1, 2, \quad (2.25)$$

$$g_{12} = \frac{\mu(t_1 - t_2)}{\delta}, \quad g_{i3} = \frac{\mu t_i}{\delta}, \quad i = 1, 2, \quad (2.26)$$

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

From (2.26), the tax sensitivity of profit shifting is $\frac{\partial(\mathbf{1}_i g_{i2} + g_{i3})}{\partial t_i} = \frac{2\mu}{\delta}$. A higher tax deductibility makes profit shifting more sensitive to tax changes and indicates a loss of revenue efficiency.

Using (2.25), we can derive the following comparative statics:

$$\frac{\partial y_i}{\partial t_i} = -\frac{\partial m_i / \partial t_i}{c''(y_i)} = -\frac{1-\mu}{(1-t_i\mu)^2 c''(y_i)} < 0, \quad \forall \mu \in [0,1), \quad (2.27)$$

$$\frac{\partial y_i}{\partial \mu} = -\frac{\partial m_i / \partial \mu}{c''(y_i)} = \frac{(1-t_i)t_i}{(1-t_i\mu)^2 c''(y_i)} > 0, \quad \forall \mu \in [0,1). \quad (2.28)$$

As in Best et al. (2015), (2.27) shows that increasing one country's tax rate raises the EMTR and reduces real output in this country. From (2.28), 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 (2.23). In doing so, it takes into account the MNE's response to countries' tax rates as in (2.25) and (2.26). 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:

$$\frac{\partial R_i(t_i, t_j)}{\partial t_i} = y - \mu c(y) - \frac{t(1-\mu)^2}{(1-\mu t)^3 c''(y)} - \frac{3\mu^2 t}{\delta} = 0, \quad (2.29)$$

where y is determined by (2.25) (with country subscripts dropped due to symmetry).

Given the specification of production cost function, solving (2.29) 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 2.A.5). Besides, (2.25) 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 2.A.6, we show

that $\forall \mu \in (0,1)$, $\tilde{y}$ lies on the interval $(\underline{y}, \bar{y})$. Given (2.23) and (2.26), each country's equilibrium revenue $\tilde{R}(\delta, \mu)$ is expressed as:

$$\tilde{R}(\delta, \mu) = \tilde{t} \tilde{\pi}^t(\tilde{t}, \delta, \mu), \quad (2.30)$$

where $\tilde{\pi}^t(\tilde{t}, \delta, \mu) := \tilde{y}(\tilde{t}, \mu) - \mu [c(\tilde{y}(\tilde{t}, \mu)) + \tilde{g}(\tilde{t}, \delta, \mu)]$ is the equilibrium tax base, with $\tilde{g}(\tilde{t}, \delta, \mu) = \frac{\mu \tilde{t}}{\delta}$ 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 $0 \leq \mu_1 < \mu_2 \leq 1$, 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 2.A.7.

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 $\delta \rightarrow 0$, the effect of tax capacity on the equilibrium revenue converges to:

$$\lim_{\delta \rightarrow 0} \frac{\partial \tilde{R}(\delta, \mu)}{\partial \delta} = \frac{2[\bar{y} - \mu c(\bar{y})]^2}{9\mu^2} \quad (\text{cf. (A27)}), \text{ which strictly decreases with deductibility}$$

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 (2.30), we decompose the effect of tax deductibility on the equilibrium tax revenue as follows:

$$\frac{\partial \tilde{R}(\delta, \mu)}{\partial \mu} = \underbrace{\frac{\partial(\tilde{t} \tilde{\pi}'(\tilde{t}, \delta, \mu))}{\partial \tilde{t}} \cdot \frac{\partial \tilde{t}(\delta, \mu)}{\partial \mu}}_{\text{indirect effect}} + \underbrace{\tilde{t} \cdot \frac{\partial \tilde{\pi}'(\tilde{t}, \delta, \mu)}{\partial \mu}}_{\text{tax base effect}}.$$

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., $\frac{\partial(\tilde{t} \tilde{\pi}'(\tilde{t}, \delta, \mu))}{\partial \tilde{t}} > 0$,

$\forall \mu \in (0, 1]$. 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. (A32) and (A33)).

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

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

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. (A33) and (A35)). Together with the positive indirect effect under these conditions, the DM theorem is extended to international profit shifting context.

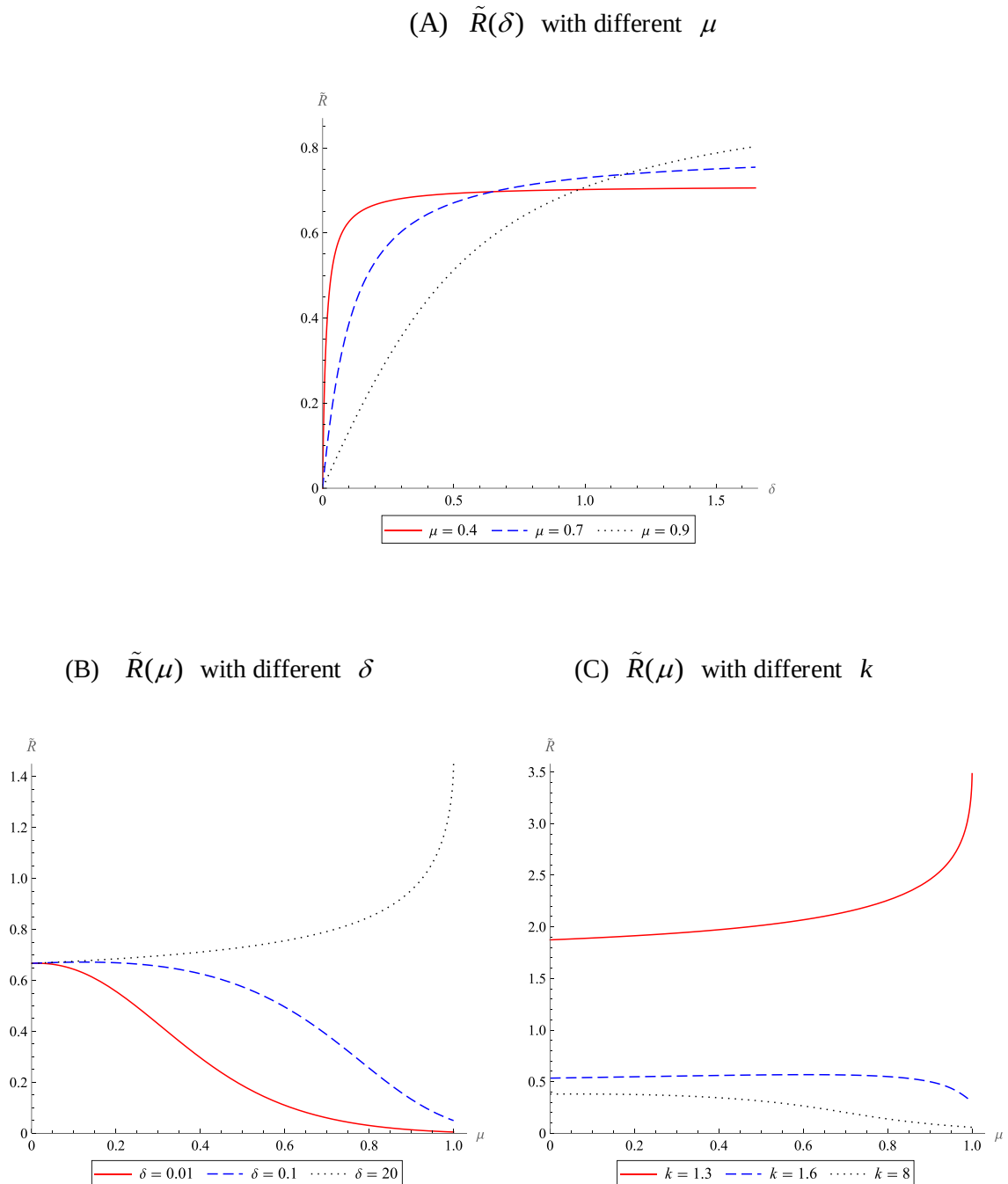


Fig. 2.2. The effects of tax deductibility on tax revenue. Panel A illustrates equilibrium tax revenues under different deductibility rates and shows how tax rankings are related to the tax capacity. Panels B

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. 2.2 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.

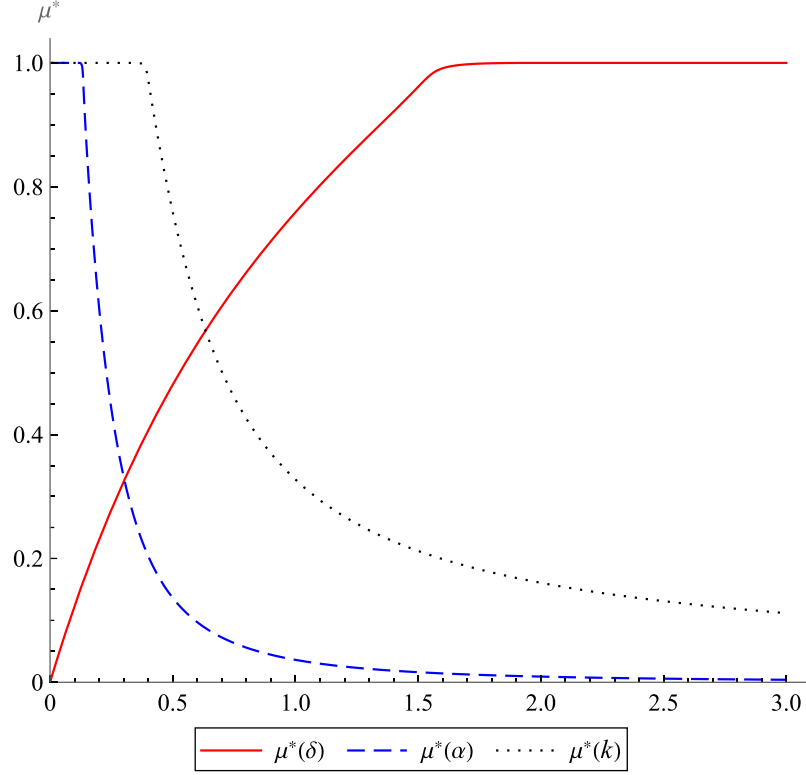


Fig. 2.3. The effects of δ , α and k on the optimal deductibility. The parameter values are specified as: $\delta = 1$, $\alpha = \frac{1}{6}$, $\beta = \frac{1}{2}$ and $k = \frac{3}{2}$. To plot each curve, we fix the values of three parameters such that the optimal deductibility is a univariate function 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. 2.3 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. 2.2. 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.

2.6. 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. By contrast, countries' equilibrium taxes strictly increase 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 across industries, depending on countries' tax capacity and industry-specific 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 2.A. Proofs.

2.A.1. Proof of Lemma 2

From (2.11), the equilibrium profit outflow from each FA country to the tax haven is given by:

$$\hat{g} = \frac{\hat{t}}{\delta}. \quad (\text{A1})$$

Inserting (A1) into (2.16) yields:

$$\bar{\pi} - \frac{3}{2}\hat{g} - \frac{\delta\hat{g}(\bar{\pi} - \hat{g})^2}{2(1 - \delta\hat{g})\bar{y}^2 c''(\bar{y})} = 0, \quad (\text{A2})$$

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

$$\delta = \frac{(2\bar{\pi} - 3\hat{g})\bar{y}^2 c''(\bar{y})}{\hat{g} \left[(\bar{\pi} - \hat{g})^2 + (2\bar{\pi} - 3\hat{g})\bar{y}^2 c''(\bar{y}) \right]}. \quad (\text{A3})$$

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

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

$$\xi'(g) = - \frac{\left[2(\bar{\pi} - g)(\bar{\pi}^2 - 3g\bar{\pi} + 3g^2) + (2\bar{\pi} - 3g)^2 \bar{y}^2 c''(\bar{y}) \right] \bar{y}^2 c''(\bar{y})}{g^2 \left[(\bar{\pi} - g)^2 + (2\bar{\pi} - 3g)\bar{y}^2 c''(\bar{y}) \right]^2} < 0. \quad (\text{A4})$$

(A4) shows that $\xi(g)$ is strictly decreasing in g over $(0, \frac{2\bar{\pi}}{3})$, which guarantees a unique $\hat{g} \in (0, \frac{2\bar{\pi}}{3})$ satisfying $\xi(\hat{g}) = \delta$. In addition, it follows from (A1) and (A2) that: $\hat{t} \in (0, 1) \Leftrightarrow \hat{g} \in (0, \frac{1}{\delta}) \Leftrightarrow \hat{g} \in (0, \frac{2\bar{\pi}}{3})$. Thus, for any $\delta > 0$, there exists a unique symmetric Nash equilibrium tax $\hat{t} \in (0, 1)$ that is determined by (2.16).

Now we prove (ii). Given (A4), it is obvious that:

$$\frac{d\hat{g}}{d\delta} = \frac{1}{\xi'(g)} < 0. \quad (\text{A5})$$

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

Using (A1) and (A3), we rewrite the equilibrium tax $\hat{t}$ as a function of $\hat{g}$:

$$\hat{t} = \delta \hat{g} = \frac{(2\bar{\pi} - 3\hat{g})\bar{y}^2 c''(\bar{y})}{(\bar{\pi} - \hat{g})^2 + (2\bar{\pi} - 3\hat{g})\bar{y}^2 c''(\bar{y})}, \quad (\text{A6})$$

from which we can derive:

$$\frac{d\hat{t}}{d\delta} = \frac{d\hat{t}}{d\hat{g}} \cdot \frac{d\hat{g}}{d\delta} = \frac{(\bar{\pi} - \hat{g})(\bar{\pi} - 3\hat{g})\bar{y}^2 c''(\bar{y})}{\left[(\bar{\pi} - \hat{g})^2 + (2\bar{\pi} - 3\hat{g})\bar{y}^2 c''(\bar{y})\right]^2} \cdot \frac{d\hat{g}}{d\delta}. \quad (\text{A7})$$

Then it follows from (A4), (A5) and (A7) that:

$$\frac{d\hat{t}}{d\delta} \begin{matrix} \geq \\ \leq \end{matrix} 0 \Leftrightarrow \hat{g} \begin{matrix} \geq \\ \leq \end{matrix} \frac{\bar{\pi}}{3} \Leftrightarrow \delta \begin{matrix} \leq \\ \geq \end{matrix} \delta^*, \quad (\text{A8})$$

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

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 (A5) – (A7), we can derive the following:

$$\frac{d\hat{R}}{d\delta} = \hat{t}\left(-\frac{d\hat{g}}{d\delta}\right) + (\bar{\pi} - \hat{g})\frac{d\hat{t}}{d\delta} = -\frac{\left[(\bar{\pi} - \hat{g})^2 \bar{\pi} + (2\bar{\pi} - 3\hat{g})^2 \bar{y}^2 c''(\bar{y})\right] \bar{y}^2 c''(\bar{y})}{\left[(\bar{\pi} - \hat{g})^2 + (2\bar{\pi} - 3\hat{g})\bar{y}^2 c''(\bar{y})\right]^2} \cdot \frac{d\hat{g}}{d\delta} > 0.$$

Q.E.D.

2.A.2. 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 (2.7), (2.22) and Lemma 1(iii), we have the following:

$$\delta^{SA} \begin{matrix} \leq \\ \geq \end{matrix} \frac{3}{\bar{\pi}} \Leftrightarrow \tilde{R}(\delta^{SA}) \begin{matrix} \leq \\ \geq \end{matrix} \tilde{R}\left(\frac{3}{\bar{\pi}}\right) \Leftrightarrow R^T \begin{matrix} \leq \\ \geq \end{matrix} \tilde{R}\left(\frac{3}{\bar{\pi}}\right) \Leftrightarrow k^{\frac{1}{1-k}} \bar{\pi} \begin{matrix} \leq \\ \geq \end{matrix} \frac{2\bar{\pi}}{3} \Leftrightarrow k \begin{matrix} \leq \\ \geq \end{matrix} 4.939, \quad (\text{A9})$$

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

Then it follows from (2.7), (2.22) and (A9) that:

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

$$\delta^{SA} = \begin{cases} \frac{9k^{\frac{1}{1-k}}}{2\bar{\pi}} & \text{if } 1 < k < 4.939 \\ \frac{1}{\bar{\pi}(1-k^{\frac{1}{1-k}})} & \text{if } k \geq 4.939 \end{cases}. \quad (\text{A10})$$

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 (A3), (2.22) and $\bar{y}^2 c''(\bar{y}) = k\bar{\pi}$, we get the following:

$$\begin{aligned} \hat{R}(\delta^{FA}) = R^T &\Leftrightarrow \frac{k\bar{\pi}(\bar{\pi} - g^{FA})(2\bar{\pi} - 3g^{FA})}{(\bar{\pi} - g^{FA})^2 + k\bar{\pi}(2\bar{\pi} - 3g^{FA})} = k^{\frac{1}{1-k}}\bar{\pi} \Leftrightarrow \\ f(g^{FA}) &:= (g^{FA})^2 \left(3k^{\frac{k}{k-1}} - 1 \right) + g^{FA} \left(2 + 3k - 5k^{\frac{k}{-1+k}} \right) \bar{\pi} + \left(2k^{\frac{k}{-1+k}} - 2k - 1 \right) \bar{\pi}^2 = 0 \Leftrightarrow \\ g^{FA} &= \frac{\left(5k^{\frac{k}{k-1}} - 3k - 2 - \sqrt{\Omega} \right) \bar{\pi}}{6k^{\frac{k}{k-1}} - 2}, \text{ where } \Omega := k^{\frac{2k}{k-1}} - 6k^{\frac{2k-1}{k-1}} + k(4+9k) > 0, \forall k > 1. \end{aligned}$$

Plugging g^{FA} into (A3) and using $\bar{y}^2 c''(\bar{y}) = k\bar{\pi}$ yields the expression for δ^{FA} :

$$\delta^{FA} = \frac{\phi(k)}{\bar{\pi}}, \quad (\text{A11})$$

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

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

$$\tilde{R}(\delta) \leq R^T \Leftrightarrow \delta \leq \delta^{SA}, \text{ and } \hat{R}(\delta) \leq R^T \Leftrightarrow \delta \leq \delta^{FA}. \quad (\text{A12})$$

Notably, when $\delta = +\infty$, it follows from (2.10) and (2.11) that $\forall (t_1, t_2) \in [0, 1]^2$, $g_{ij} = 0$ 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 (A10) lead to the effects of α and k on threshold δ^{SA} .

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

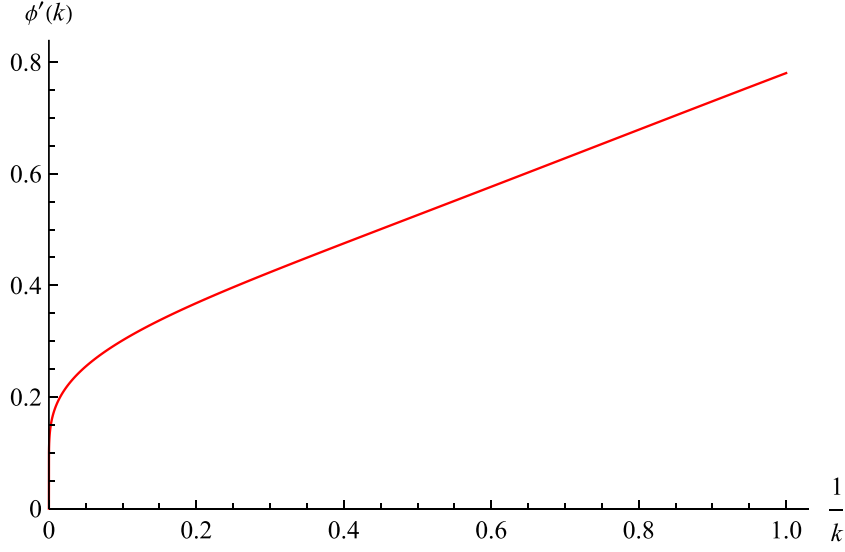


Fig. A1. The graph of $\phi'(k)$.

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 (A11), the comparative static effects on δ^{FA} immediately follow. **Q.E.D.**

2.A.3. 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 (2.6), (2.7), (A1) and (A8), we can derive the following:

$$\tilde{R} \begin{matrix} \leq \\ \geq \end{matrix} \hat{R} \Leftrightarrow \frac{2\bar{\pi}^2\delta}{9} \begin{matrix} \leq \\ \geq \end{matrix} \delta\hat{g}(\bar{\pi} - \hat{g}) \Leftrightarrow \hat{g} \begin{matrix} \geq \\ \leq \end{matrix} \frac{\bar{\pi}}{3} \Leftrightarrow \hat{t} \begin{matrix} \geq \\ \leq \end{matrix} \tilde{t} \Leftrightarrow \delta \begin{matrix} \leq \\ \geq \end{matrix} \delta^*, \quad (\text{A13})$$

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

Case (ii): $\delta \geq \frac{3}{\bar{\pi}}$. It follows from (2.6) and Lemma 2(i) that:

$$\tilde{t} = 1 > \hat{t} > 0. \quad (\text{A14})$$

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

$$\forall \delta \geq \frac{3}{\pi}, \quad \tilde{R} = \tilde{t}(\bar{\pi} - \frac{\tilde{t}}{\delta}) > \hat{t}(\bar{\pi} - \frac{\hat{t}}{\delta}) = \hat{R}. \quad (\text{A15})$$

Combining (A13) – (A15) leads to Lemma 4. **Q.E.D.**

2.A.4. Proof of Proposition 2

We first rank δ^{SA} , δ^{FA} and δ^* , conditional on the value of k . Plugging $\delta = \delta^*$ into (2.7) and using (2.22) leads to:

$$\tilde{R}(\delta^*) \underset{\leq}{\geq} R^T \Leftrightarrow \frac{6k\bar{\pi}}{4+9k} \underset{\leq}{\geq} k^{\frac{1}{1-k}} \bar{\pi} \Leftrightarrow k \underset{\leq}{\geq} 3.089, \quad (\text{A16})$$

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 (A12) and Lemma 4, we can derive:

$$\tilde{R}(\delta^*) \underset{\leq}{\geq} R^T \Rightarrow \delta^* \underset{\leq}{\geq} \delta^{SA} \Rightarrow \hat{R}(\delta^{SA}) \underset{\leq}{\geq} \tilde{R}(\delta^{SA}) = R^T \Rightarrow \delta^{SA} \underset{\leq}{\geq} \delta^{FA}. \quad (\text{A17})$$

Combining (A16) and (A17) yields the ranking of the three thresholds:

$$\delta^* \underset{\leq}{\geq} \delta^{SA} \underset{\leq}{\geq} \delta^{FA} \Leftrightarrow k \underset{\leq}{\geq} 3.089. \quad (\text{A18})$$

Given Proposition 1(i), Lemma 4 and (A18), 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 (A18), it is straightforward to show the following:

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

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

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

2.A.5. The existence and uniqueness of symmetric equilibrium under SA

As two polar cases (i.e., $\mu \in \{0, 1\}$) have been studied in Section 2.3, we here focus on

the case of partial deductibility with $\mu \in (0,1)$. Given the specific functional form of $c(\cdot)$, it follows from (2.25) that $\tilde{t} \leq \frac{1-\beta}{1-\beta\mu} \leq 1$. Define

$F(t, \mu, \delta) := y - \mu c(y) - \frac{t(1-\mu)^2}{(1-\mu t)^3 c''(y)} - \frac{3\mu^2 t}{\delta}$, where y is determined by (2.25). It is obvious that: $\forall \mu \in (0,1)$, $F(0, \mu, \delta) = \bar{y} - \mu c(\bar{y}) > 0$ and $\lim_{t \rightarrow (1-\beta)/(1-\beta\mu)^-} F(t, \mu, \delta) < 0$. This property together with the continuity of $F(t, \mu, \delta)$ in t ensures the existence of a Nash equilibrium with tax rate $\tilde{t} \in \left(0, \frac{1-\beta}{1-\beta\mu}\right)$.

On the other hand, solving for $\tilde{y}$ from (2.25) yields $\tilde{y} = \left[\frac{1-\tilde{t} - \beta(1-\mu\tilde{t})}{\alpha(1-\mu\tilde{t})} \right]^{\frac{1}{k-1}}$. Then plugging this expression into (2.29) and rearranging terms gives:

$$H(k, \beta, \mu, \tilde{t}) = \frac{3\mu^2 \tilde{t}}{\delta} \cdot \alpha^{\frac{1}{k-1}} \left(\frac{1-\tilde{t}}{1-\mu\tilde{t}} - \beta \right)^{\frac{k-2}{k-1}} > 0, \quad (\text{A19})$$

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

Given (2.25), (2.27) and the specification of $c(\cdot)$, computing $\frac{\partial F(t, \mu, \delta)}{\partial t}$ at Nash equilibrium tax rate $\tilde{t}$ and simplifying yields:

$$\frac{\partial F(\tilde{t}, \mu, \delta)}{\partial \tilde{t}} = - \frac{(1-\mu)^2 \Phi_1}{(k-1)(1-\mu\tilde{t})^4 [1-\tilde{t} - \beta(1-\mu\tilde{t})] c''(\tilde{y})} - \frac{3\mu^2}{\delta}, \quad (\text{A20})$$

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

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

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

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

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

From the above contradiction, it must be that $\Phi_1 > 0$. Then it follows from (A20) that:

$$\forall \mu \in (0,1), \quad \frac{\partial F(\tilde{t}, \mu, \delta)}{\partial \tilde{t}} < 0. \quad (\text{A21})$$

(A21) 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.

2.A.6. The range of the equilibrium output under SA

In Section 2.3, we have obtained each country's equilibrium output under turnover and profit taxation, i.e., $\tilde{y} = \underline{y}$ if $\mu = 0$, and $\tilde{y} = \bar{y}$ if $\mu = 1$. Here we investigate the partial deductibility with $\mu \in (0,1)$. Given (2.25) 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 $\tilde{y} < \bar{y}$. On the other hand, it follows from (2.25) that $\tilde{t} = \frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}$. Inserting this expression into (A19) yields:

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

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 $\tilde{y} \leq \underline{y}$, then we would derive the following:

$$\begin{aligned} \tilde{y} \leq \underline{y} &\Rightarrow c'(\tilde{y}) \leq c'(\underline{y}) = \frac{1-\beta}{k} + \beta \Rightarrow k \leq \frac{1-\beta}{c'(\tilde{y})-\beta} \Rightarrow \\ 0 &< H\left(k, \beta, \mu, \frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}\right) \leq H\left(\frac{1-\beta}{c'(\tilde{y})-\beta}, \beta, \mu, \frac{1-c'(\tilde{y})}{1-\mu c'(\tilde{y})}\right) = -\frac{\mu(1-\mu\beta)[c'(\tilde{y})-\beta][1-c'(\tilde{y})]^2}{(1-\mu)(1-\beta)} < 0, \end{aligned}$$

which is a contradiction. Therefore, we must have:

$$\underline{y} < \tilde{y} < \bar{y}, \quad \forall \mu \in (0,1). \quad (\text{A22})$$

2.A.7. Proof of Proposition 3

Given (2.25), (2.27), (2.29) and (2.30), the effect of tax capacity on each country's equilibrium revenue is:

$$\frac{\partial \tilde{R}(\delta, \mu)}{\partial \delta} = \frac{\partial(\tilde{t} \tilde{\pi}'(\tilde{t}, \delta, \mu))}{\partial \tilde{t}} \cdot \frac{\partial \tilde{t}(\delta, \mu)}{\partial \delta} + \tilde{t} \cdot \frac{\partial \tilde{\pi}'(\tilde{t}, \delta, \mu)}{\partial \delta}, \quad (\text{A23})$$

where

$$\frac{\partial(\tilde{t}\tilde{\pi}'(\tilde{t},\delta,\mu))}{\partial\tilde{t}} = \tilde{y} - \mu c(\tilde{y}) - \frac{\mu^2\tilde{t}}{\delta} + \tilde{t} \left[(1 - \mu c'(\tilde{y})) \frac{\partial\tilde{y}(\tilde{t},\mu)}{\partial\tilde{t}} - \frac{\mu^2}{\delta} \right] = \frac{\mu^2\tilde{t}}{\delta} > 0, \quad (\text{A24})$$

$$\frac{\partial\tilde{\pi}'(\tilde{t},\delta,\mu)}{\partial\delta} = \frac{\mu^2\tilde{t}}{\delta^2} > 0. \quad (\text{A25})$$

On the other hand, using (2.25) and (2.29), we can derive the following:

$$\forall \mu \in (0,1], \lim_{\delta \rightarrow 0} \tilde{t}(\delta, \mu) = 0, \lim_{\delta \rightarrow 0} \tilde{y}(\tilde{t}, \mu) = \bar{y}, \text{ and } \lim_{\delta \rightarrow 0} \frac{\tilde{t}(\delta, \mu)}{\delta} = \lim_{\delta \rightarrow 0} \frac{\partial\tilde{t}(\delta, \mu)}{\partial\delta} = \frac{\bar{y} - \mu c(\bar{y})}{3\mu^2}, \quad (\text{A26})$$

where the first equality in the third part of (A26) holds by L'Hôpital's rule.

Taking the limit on both sides of (A23) as $\delta \rightarrow 0$ and using (A24) – (A26) yields:

$$\forall \mu \in (0,1], \lim_{\delta \rightarrow 0} \frac{\partial\tilde{R}(\delta, \mu)}{\partial\delta} = \lim_{\delta \rightarrow 0} \left(\frac{\mu^2\tilde{t}}{\delta} \cdot \frac{\partial\tilde{t}(\delta, \mu)}{\partial\delta} + \frac{\mu^2\tilde{t}^2}{\delta^2} \right) = \frac{2[\bar{y} - \mu c(\bar{y})]^2}{9\mu^2}. \quad (\text{A27})$$

(A27) indicates that $\lim_{\delta \rightarrow 0} \frac{\partial\tilde{R}(\delta, \mu)}{\partial\delta}$ is strictly decreasing in μ , which leads to the following:

$$\begin{aligned} &\forall \mu_1, \mu_2 \in (0,1] \text{ with } \mu_1 < \mu_2, \text{ there exists some } \bar{\delta} > 0 \text{ such that} \\ &\frac{\partial}{\partial\delta} [\tilde{R}(\delta, \mu_1) - \tilde{R}(\delta, \mu_2)] > 0 \text{ when } 0 < \delta < \bar{\delta} \Rightarrow \\ &\tilde{R}(\delta, \mu_1) - \tilde{R}(\delta, \mu_2) > \lim_{\delta \rightarrow 0} [\tilde{R}(\delta, \mu_1) - \tilde{R}(\delta, \mu_2)] = 0 \text{ when } 0 < \delta < \bar{\delta}. \end{aligned}$$

For the special case where $\mu_1 = 0 < \mu_2 \leq 1$, we have: $\tilde{R}(\delta, 0) = R^T > 0 = \lim_{\delta \rightarrow 0} \tilde{R}(\delta, \mu_2)$, where R^T is independent of tax capacity $\Rightarrow$ there exists some $\bar{\delta} > 0$ such that $\tilde{R}(\delta, 0) > \tilde{R}(\delta, \mu_2)$ when $0 < \delta < \bar{\delta}$.

Now we prove (ii). For simplicity, in what follows we assume that $\beta = 0$, while our result still holds for $0 < \beta < 1$. Given (2.30), the effect of the deductibility μ on each country's equilibrium tax revenue under SA is:

$$\frac{\partial\tilde{R}(\delta, \mu)}{\partial\mu} = \underbrace{\frac{\partial(\tilde{t}\tilde{\pi}'(\tilde{t},\delta,\mu))}{\partial\tilde{t}} \cdot \frac{\partial\tilde{t}(\delta, \mu)}{\partial\mu}}_{\text{indirect effect}} + \underbrace{\tilde{t} \cdot \frac{\partial\tilde{\pi}'(\tilde{t},\delta,\mu)}{\partial\mu}}_{\text{tax base effect}}. \quad (\text{A28})$$

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 (2.25), (2.28), (2.29) and the production cost function specification,

differentiating $F(\tilde{t}, \mu, \delta)$ w.r.t. μ and then heavily rearranging terms gives:

$$\frac{\partial F(\tilde{t}, \mu, \delta)}{\partial \mu} = \frac{(1-\mu\tilde{t})}{1-\mu} \left[\tilde{y}c'(\tilde{y}) - c(\tilde{y}) - \frac{3\mu\tilde{t}(2-\mu-\mu\tilde{t})}{(1-\mu\tilde{t})\delta} \right] + \frac{\tilde{t}(1-\mu)}{(1-\mu\tilde{t})^4 (k-1)c''(\tilde{y})} [2k(1-\tilde{t}) + \tilde{t}(\mu+1) - 2]. \quad (\text{A29})$$

Notice that $\tilde{y}c'(\tilde{y}) - c(\tilde{y})$ is strictly increasing in $\tilde{y}$, and recall from (A22) 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 (A29), we have:

$$\tilde{y}c'(\tilde{y}) - c(\tilde{y}) - \frac{3\mu\tilde{t}(2-\mu-\mu\tilde{t})}{(1-\mu\tilde{t})\delta} > 0 \quad \text{if} \quad \delta \geq \frac{6}{\underline{y}c'(\underline{y}) - c(\underline{y})}. \quad (\text{A30})$$

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: $k \leq \frac{2-\tilde{t}(\mu+1)}{2(1-\tilde{t})} \Rightarrow$

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

Therefore, given (A29) and (A30) we get:

$$\forall \mu \in (0, 1), \quad \frac{\partial F(\tilde{t}, \mu, \delta)}{\partial \mu} > 0 \quad \text{if} \quad \delta \geq \frac{6}{\underline{y}c'(\underline{y}) - c(\underline{y})}. \quad (\text{A31})$$

Given $F(\tilde{t}, \mu, \delta) = 0$, using (A21), (A31) and the implicit function theorem yields:

$$\frac{\partial \tilde{t}(\mu, \delta)}{\partial \mu} = -\frac{\partial F(\tilde{t}, \mu, \delta) / \partial \mu}{\partial F(\tilde{t}, \mu, \delta) / \partial \tilde{t}} > 0 \quad \text{if} \quad \delta \geq \frac{6}{\underline{y}c'(\underline{y}) - c(\underline{y})}. \quad (\text{A32})$$

Moreover, using (2.21), 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 $\lim_{k \rightarrow 1^+} \varphi(\alpha, k) = +\infty$ and $\lim_{k \rightarrow +\infty} \varphi(\alpha, k) = 0$. Then we can derive that:

$$\delta \geq \frac{6}{\underline{y}c'(\underline{y}) - c(\underline{y})} \Leftrightarrow \alpha \leq \left[\frac{\delta(k-1)}{6} \right]^{k-1} k^{1-2k} \Leftrightarrow k \leq \bar{k}, \quad (\text{A33})$$

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

Given (A24), (A32) and (A33), the indirect effect of tax deductibility is positive for a high tax capacity, low marginal production cost or large output elasticity.

Step 2:

Given (2.25) and (2.28) – (2.30), computing the tax base effect and then rearranging terms gives:

$$\begin{aligned} \tilde{t} \cdot \frac{\partial \tilde{\pi}^t(\tilde{t}, \delta, \mu)}{\partial \mu} &= \tilde{t} \left[\underbrace{(1 - \mu c'(\tilde{y})) \frac{\partial \tilde{y}(\tilde{t}, \mu)}{\partial \mu}}_{\text{production effect}} - \underbrace{(c(\tilde{y}) + \tilde{g})}_{\text{deduction effect}} - \underbrace{\mu \frac{\partial \tilde{g}(\tilde{t}, \delta, \mu)}{\partial \mu}}_{\text{shifting effect}} \right] \\ &= \frac{\tilde{t}(1 - \mu \tilde{t})}{1 - \mu} \left[\tilde{y} c'(\tilde{y}) - c(\tilde{y}) - \frac{\mu \tilde{t}(2 + \mu - 3\mu \tilde{t})}{(1 - \mu \tilde{t})\delta} \right]. \end{aligned} \quad (\text{A34})$$

Note that $\frac{2 + \mu - 3\mu \tilde{t}}{1 - \mu \tilde{t}} = 3 - \frac{1 - \mu}{1 - \mu \tilde{t}} \in (2, 3)$, and recall from (A22) that $\tilde{y} > \underline{y}$. Then it immediately follows from (A34) that:

$$\forall \mu \in (0, 1), \quad \tilde{t} \cdot \frac{\partial \tilde{\pi}^t(\tilde{t}, \delta, \mu)}{\partial \mu} > 0 \quad \text{if} \quad \delta \geq \frac{3}{\underline{y} c'(\underline{y}) - c(\underline{y})}. \quad (\text{A35})$$

Combining (A24), (A28), (A32) and (A35) leads to:

$$\forall \mu \in (0, 1), \quad \frac{\partial \tilde{R}(\delta, \mu)}{\partial \mu} > 0 \quad \text{if} \quad \delta \geq \frac{6}{\underline{y} c'(\underline{y}) - c(\underline{y})}. \quad (\text{A36})$$

Given (A33), (A36) and the continuity of $\tilde{R}(\delta, \mu)$ at $\mu = 0, 1$, Proposition 3(ii) directly follows. **Q.E.D.**

Appendix 2.B. The impact of tax deductibility under FA

Suppose as in Subsection 2.3.2 that the two countries form a FA union, while the tax haven sticks to SA. Using (2.23) yields the MNE's consolidated tax base within the FA union:

$$\pi^c := \sum_{i=1}^2 [y_i - \mu(c(y_i) + \mathbf{1}_i g_{12} + g_{i3})] = \sum_{i=1}^2 [y_i - \mu(c(y_i) + g_{i3})]. \quad (\text{B1})$$

The tax base of the MNE in country i is characterized by (2.9), where π^c is given by (B1). Using (2.1), (2.9) and (B1), the MNE's total after-tax profit under FA amounts to:

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

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

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

$$g_{12} = 0, \quad (B2)$$

$$g_{i3} = \frac{\mu\tau}{\delta}, \quad i = 1, 2, \quad (B3)$$

$$c'(y_i) = \frac{1-\tau}{1-\mu\tau} + \frac{(\tau-t_i)\pi^c}{(1-\mu\tau)\sum_{i=1}^2 y_i}, \quad i = 1, 2. \quad (B4)$$

As shown by (B2), the statement that profit shifting within the FA union can be eliminated still holds. On the other hand, (B3) 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 (B4) that the real output is equal in the two countries, which is still determined by (2.25) (with country subscripts dropped due to symmetry).

Totally differentiating (B3) and (B4), together with the symmetry property, yields the comparative statics under identical tax rates $t_i = t$ ($i = 1, 2$):

$$\frac{\partial g_{i3}}{\partial t_i} = \frac{\partial g_{j3}}{\partial t_i} = \frac{\mu}{2\delta} > 0, \quad i, j = 1, 2, \quad (B5)$$

$$\frac{\partial y_i}{\partial t_i} = -\frac{y[1-\mu c'(y)] + [y-\mu c(y)-\mu^2 t/\delta]}{2y(1-\mu t)c''(y)} < 0, \quad i = 1, 2, \quad (B6)$$

$$\frac{\partial y_j}{\partial t_i} = \frac{\mu[y c'(y) - c(y) - \mu t/\delta]}{2y(1-\mu t)c''(y)}, \quad i, j = 1, 2; \quad i \neq j. \quad (B7)$$

Using (2.25), (B6) and (B7), under equal tax rates $t_i = t$ ($i = 1, 2$), we can derive:

$$\forall \mu \in [0, 1), \quad \frac{\partial y_i}{\partial t_i} - \frac{\partial y_j}{\partial t_i} = -\frac{y - \mu c(y) - \mu^2 t/\delta}{y(1-\mu t)c''(y)} < 0 \quad \text{and} \quad \frac{\partial y_i}{\partial t_i} + \frac{\partial y_j}{\partial t_i} = -\frac{1-\mu}{(1-\mu t)^2 c''(y)} < 0, \quad$$

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: $\forall \mu \in [0,1]$,

$$\frac{\partial \lambda_i}{\partial t_i} = -\frac{\partial \lambda_j}{\partial t_i} = -\frac{y - \mu c(y) - \mu^2 t / \delta}{4y^2(1 - \mu t)c''(y)} < 0. \text{ As discussed in Subsection 2.3.2, the MNE}$$

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 (B1). In doing so, it takes into account the MNE's behavior that depends on countries' tax rates due to (B2) – (B4). Solving the first-order conditions of revenue maximization problem and using the symmetry property yields the following:

$$\frac{\partial R_i}{\partial t_i} = y - \mu c(y) - \frac{3\mu^2 t}{2\delta} - \frac{t \left[y - \mu c(y) - \frac{\mu^2 t}{\delta} \right]^2}{2y^2(1 - \mu t)c''(y)} - \frac{t(1 - \mu)^2}{2(1 - \mu t)^3 c''(y)} = 0, \quad (\text{B8})$$

where y is determined by (2.25) (with country subscripts dropped due to symmetry).

Given the specific functional form of $c(y)$, solving (B8) yields each FA country's Nash equilibrium tax rate $\hat{t}(\delta, \mu)$. Similar to SA, (2.25) 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 (2.9), (B1) and (B3), each FA country's equilibrium revenue is still characterized by (2.30), 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 $0 \leq \mu_1 < \mu_2 \leq 1$, 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 (2.25), (2.27) and (2.30) (with the tildes replaced by hats), we can compute the effect of tax capacity on the equilibrium tax revenue under FA:

$$\frac{\partial \hat{R}(\delta, \mu)}{\partial \delta} = \frac{\partial(\hat{t} \hat{\pi}'(\hat{t}, \delta, \mu))}{\partial \hat{t}} \cdot \frac{\partial \hat{t}(\delta, \mu)}{\partial \delta} + \hat{t} \cdot \frac{\partial \hat{\pi}'(\hat{t}, \delta, \mu)}{\partial \delta}, \quad (\text{B9})$$

where

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

Given (2.25) and (B8), we can obtain the following:

$$\begin{aligned} \forall \mu \in (0,1] \quad , \quad \lim_{\delta \rightarrow 0} \hat{t}(\delta, \mu) = 0 \quad , \quad \lim_{\delta \rightarrow 0} \hat{y}(\delta, \mu) = \bar{y} \quad , \quad \text{and} \\ \lim_{\delta \rightarrow 0} \frac{\hat{t}(\delta, \mu)}{\delta} = \lim_{\delta \rightarrow 0} \frac{\partial \hat{t}(\delta, \mu)}{\partial \delta} = \frac{2[\bar{y} - \mu c(\bar{y})]}{3\mu^2}, \end{aligned} \quad (\text{B10})$$

where the first equality in the third part of (B10) holds by L'Hôpital's rule.

Taking the limit on both sides of (B9) as $\delta \rightarrow 0$ and using (B10) yields:

$$\begin{aligned} \forall \mu \in (0,1], \quad \lim_{\delta \rightarrow 0} \frac{\partial(\hat{t} \hat{\pi}'(\hat{t}, \delta, \mu))}{\partial \hat{t}} = -\frac{\bar{y} - \mu c(\bar{y})}{3} < 0, \quad \lim_{\delta \rightarrow 0} \left(\hat{t} \cdot \frac{\partial \hat{\pi}'(\hat{t}, \delta, \mu)}{\partial \delta} \right) = \frac{4[\bar{y} - \mu c(\bar{y})]^2}{9\mu^2}, \quad \text{and} \\ \lim_{\delta \rightarrow 0} \frac{\partial \hat{R}(\delta, \mu)}{\partial \delta} = \frac{2[\bar{y} - \mu c(\bar{y})]^2}{9\mu^2}. \end{aligned} \quad (\text{B11})$$

(A27) and the third part of (B11) indicate that as $\delta \rightarrow 0$, 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, (A24) indicates that the fiscal externality under SA is positive and causes inefficient undertaxation. In sharp contrast, the first part of (B11) 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 2.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) $0 \leq c'(0) < 1$, $c'(+\infty) > 1$; (iii) $c''(y) > 0$, $2c''(y) + yc'''(y) > 0$, $\forall y > 0$;
- (iv) $\text{sign}\{c'''(y)\} = \text{constant}$, $\forall y > 0$. (C1)

Assumption (i) in (C1) 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 $c'''(.) \leq 0$, the turnover tax can cause a large loss of production efficiency, as will be shown by (C7). In the case of $c'''(.) > 0$, the value of $c''(\bar{y})$ 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 function specified in the main text satisfies all the four assumptions in (C1).

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:*

$$t^T = \underline{y}c''(\underline{y}) \text{ and } R^T = \underline{y}[1 - c'(\underline{y})] = \underline{y}^2c''(\underline{y}), \quad (\text{C2})$$

where equilibrium output $\underline{y}$ is implicitly determined by $c'(\underline{y}) + \underline{y}c''(\underline{y}) = 1$.

Proof. Using (2.19), 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:

$$c'(\underline{y}) + \underline{y}c''(\underline{y}) = 1. \quad (\text{C3})$$

Given (2.19) and (C3), (C2) immediately follows. **Q.E.D.**

Recall from (2.4) and (2.12) that the efficient output level of each country under profit tax is characterized by $c'(\bar{y})=1$. In (C3), $\underline{y}c''(\underline{y}) > 0$ reflects the “production wedge”, which makes the output under turnover tax smaller than under profit tax, i.e., $\underline{y} < \bar{y}$.

Notice that under profit tax with SA the Nash equilibrium tax rate and revenue of the two countries are still characterized by (2.6) and (2.7), respectively. Under profit tax with FA, the Nash equilibrium tax of the FA countries is still determined by (2.16). 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) $\lim_{\delta \rightarrow 0} \tilde{R}(\delta) < R^T$ and $\lim_{\delta \rightarrow 0} \hat{R}(\delta) < R^T$; (ii) $\lim_{\delta \rightarrow +\infty} \tilde{R}(\delta) > R^T$ and $\lim_{\delta \rightarrow +\infty} \hat{R}(\delta) > R^T$.*

Proof. It is obvious from (2.7) and (C2) that: $\lim_{\delta \rightarrow 0} R^{SA}(\delta) = 0 < R^T$. As for FA, it follows from (A2) that $\lim_{\delta \rightarrow 0} \hat{g} = \frac{2\bar{\pi}}{3}$. Then we can get: $\lim_{\delta \rightarrow 0} \hat{R}(\delta) = \lim_{\delta \rightarrow 0} (\delta \hat{g})(\bar{\pi} - \hat{g}) = 0 < R^T$.

We now prove (ii). Given (2.7) and (C2), it is straightforward to show:

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

On the other hand, taking the limit on both sides of (2.16) as $\delta \rightarrow +\infty$, we can derive:

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

$$\lim_{\delta \rightarrow +\infty} \hat{t} = \frac{2\bar{y}^2 c''(\bar{y})}{\bar{\pi} + 2\bar{y}^2 c''(\bar{y})} \Rightarrow \lim_{\delta \rightarrow +\infty} \hat{R}(\delta) = \lim_{\delta \rightarrow +\infty} \hat{t}(\bar{\pi} - \frac{\hat{t}}{\delta}) = \frac{2\bar{\pi}\bar{y}^2 c''(\bar{y})}{\bar{\pi} + 2\bar{y}^2 c''(\bar{y})}. \quad (C4)$$

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

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

From the second part of assumption (iii) in (C1), $y^2 c''(y)$ is strictly increasing in y . Then it follows from $\bar{y} > \underline{y}$ that:

$$\bar{y}^2 c''(\bar{y}) > \underline{y}^2 c''(\underline{y}) = R^T. \quad (C5)$$

Besides, notice that:

$$\forall y > 0, 1 - c'(y) \geq -c''(\underline{y})(y - \underline{y}) + 1 - c'(\underline{y}) = (2\underline{y} - y)c''(\underline{y}), \quad (C6)$$

where the first inequality derives from the concavity of function $c'(\cdot)$, and the last equality holds due to (C3).

Letting $y = \bar{y}$ in (C6), we get: $1 - c'(\bar{y}) = 0 \geq (2\underline{y} - \bar{y})c''(\underline{y})$, which leads to:

$$\bar{y} \geq 2\underline{y}. \quad (C7)$$

Given (C2), (C6) and (C7), it is straightforward to show:

$$\bar{\pi} = \int_0^{\bar{y}} [1 - c'(y)] dy \geq \int_0^{2\underline{y}} [1 - c'(y)] dy \geq \int_0^{2\underline{y}} (2\underline{y} - y)c''(\underline{y}) dy = 2\underline{y}^2 c''(\underline{y}) = 2R^T. \quad (C8)$$

Combining (C4), (C5) and (C8) yields:

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

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

Firstly, notice that:

$$\forall \theta \in (0, 1), \quad 1 - c'((1 - \theta)\underline{y} + \theta\bar{y}) > 1 - (1 - \theta)c'(\underline{y}) - \theta c'(\bar{y}) = (1 - \theta)[1 - c'(\underline{y})], \quad (C9)$$

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

Then we can show the following:

$$\begin{aligned} \bar{\pi} &= \int_0^{\underline{y}} [1 - c'(y)] dy + \int_{\underline{y}}^{\bar{y}} [1 - c'(y)] dy \\ &= \int_0^{\underline{y}} [1 - c'(y)] dy + \int_0^1 (\bar{y} - \underline{y}) [1 - c'((1 - \theta)\underline{y} + \theta\bar{y})] d\theta \\ &> \int_0^{\underline{y}} [1 - c'(\underline{y})] dy + \int_0^1 (\bar{y} - \underline{y})(1 - \theta)[1 - c'(\underline{y})] d\theta \\ &= \frac{\underline{y} + \bar{y}}{2} [1 - c'(\underline{y})] = \frac{1}{2} (\underline{y} + \bar{y}) \underline{y} c''(\underline{y}) \end{aligned} \quad (C10)$$

where the first inequality follows from $c''(\cdot) > 0$ and (C9), and the last equality is obtained by using (C3).

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

$$\bar{y}^2 c''(\bar{y}) > \bar{y}^2 \frac{c'(\bar{y}) - c'(\underline{y})}{\bar{y} - \underline{y}} = \frac{\bar{y}^2 \underline{y} c''(\underline{y})}{\bar{y} - \underline{y}}, \quad (C11)$$

where the last equality is obtained by $c'(\bar{y}) = 1$ and (C3).

Combining (C4), (C10) and (C11) leads to:

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

from which we have:

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

Therefore, in both cases, $\lim_{\delta \rightarrow +\infty} \hat{R}(\delta) > R^T$ 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 for the general form of production technology.

Chapter 3. The Global Minimum Tax, Investment Incentives and Asymmetric Tax Competition

Abstract

This paper investigates how the OECD's global minimum tax (GMT) affects multinational enterprises (MNEs) behavior and countries' corporate taxes. We consider both profit shifting and capital investment responses of the MNE in a formal model of tax competition between asymmetric countries. The GMT reduces the true tax rate differential and benefits the large country, while the revenue effect is generally ambiguous for the small country. In the short run where tax rates are fixed, due to tax deduction of the substance-based income exclusion (SBIE), a higher minimum rate exerts investment incentives but also incurs a larger revenue loss for the small country. We show that under high (low) profit shifting costs the former (latter) effect dominates so that the small country's revenue increases (decreases). In the long run where countries can adjust tax rates, the GMT reshapes the tax game and the competition pattern. In contrast to the existing literature, we reveal that the minimum rate binds the small country only if it is low. With the rise of the GMT rate, countries will undercut the minimum to boost real investments and collect top-up taxes. Besides, a marginal GMT reform does not necessarily benefit the small country. It can raise the small country's tax revenue under large market-size asymmetry or low profit shifting cost, and may harm the small country otherwise.

JEL classification: F21, F23, H25, H73, H87

Keywords: Corporate taxes, Global minimum tax, Profit shifting, SBIE, Tax competition

3.1. Introduction

Multinational enterprises (MNEs) can exploit loopholes in tax rules to shift profits to low-tax countries to avoid paying taxes. International profit shifting has caused considerable losses of tax revenue for both OECD and developing countries (see Crivelli, de Mooij and Keen, 2016; Davies et al., 2018; Bilicka, 2019; Wier and Zucman, 2022; Tørsløv et al., 2023). To relieve the pressure on the outdated international corporate tax system, over 135 jurisdictions in 2021 agreed on a 15% global minimum tax (GMT), which is a key part of Pillar Two of the two-pillar solution proposed by the OECD/G20 Inclusive Framework on base erosion and profit shifting (BEPS). The GMT ensures that large MNEs with revenues above €750 million are subject to a 15% minimum tax rate in each jurisdiction where they operate.

It is expected that around 90% of in-scope MNEs will be subject to the GMT by 2025, based on the jurisdictions that have implemented or announced implementation (see Hugger et al., 2024). In practice, the GMT under Pillar Two works in a roundabout way. The first step is to determine the jurisdictional effective tax rate (ETR) of the MNE. It is computed by dividing the taxes (called covered taxes) of the affiliate in that jurisdiction by the income (called GloBE income) it has. If the jurisdictional ETR is below 15%, the MNE will be subject to a top-up tax. The top-up tax rate is the difference between the 15% minimum rate and the ETR. The top-up tax base (called excess profit) is calculated as the GloBE income in excess of the Substance-Based Income Exclusion (SBIE). The SBIE allows MNEs to tax-deduct a percentage of the carrying value of tangible assets and payroll expenses from the GloBE income of the low-tax affiliate. The current carve-out rate is 8% on tangible assets and 10% on payroll costs, while it will reduce to 5% on both tangibles and payroll over a transition period of ten years (see OECD, 2021; European Commission, 2021; Devereux et al., 2022). Hence, the SBIE benefits the low-tax affiliates that have real economic activity by reducing their top-up tax liability. A key question is which countries receive the additional tax revenue. In principle, the top-up tax can be collected either by the country where the headquarters of the MNE resides or by the host country where the MNE's affiliate records the profit. The first scenario corresponds to the income inclusion rule (IIR), while the latter corresponds to the qualified domestic minimum top-up tax (QDMTT). The Model Rules published by the OECD/G20 Inclusive Framework in December 2021 introduced the QDMTT, which gives the host country the priority to collect top-up taxes over the headquarters country (see OECD, 2021). While jurisdictions are not required to implement a QDMTT, there is a very strong incentive for countries affected by Pillar Two to do so. Notice that an MNE's tax liability is the same, no matter which country collects the top-up tax. So failure to adopt the QDMTT will cede tax revenue to other countries while conveying no tax benefit to the MNE. See IMF (2023) for detailed explanations of why countries should adopt the QDMTT.

In this paper, we investigate the OECD's GMT in a formal model of international tax competition with two *asymmetric* countries that differ in market size. First, two countries choose corporate tax rates noncooperatively to maximize their revenues. Then the MNE chooses capital investment in each country and profit shifting level to maximize total after-tax profits, taking as given the tax environments. Following the previous literature on minimum taxation (see Kanbur and Keen, 1993; Keen and Konrad, 2013), we focus on the situation where the minimum rate lies between two countries' initial tax rates without the GMT. Indeed, choosing a minimum that is above the higher of the initial tax rates is unappealing to policymakers of both countries and is politically infeasible. Our model captures the key features of Pillar Two (i.e., the SBIE and the QDMTT), and takes into account both profit shifting and real investment responses of the MNE. We analyze how the GMT affects the MNE's behavior and countries' taxes from a short-run perspective where corporate tax rates are fixed, and from a long-run perspective where governments engage in tax competition. This approach disentangles the tax incentive effect of the GMT due to the tax deduction of the SBIE from the behavioral adjustments by the governments. Similar treatment is used in the literature that compares two alternative tax principles – separate accounting (SA) and formula apportionment (FA) – in the taxation of MNEs (e.g., Riedel and Runkel, 2007; Mardan and Stimmelmayer, 2018). Different from SA under which profit is taxed in the country where the MNE declares it, under FA the MNE's taxable incomes are consolidated first and then assigned to each country based on a formula. Notably, Pillar One reallocates 25% of large MNEs' residual profits – which are the profits in excess of 10% of the revenues – to market jurisdictions through a sales-based formula (see OECD, 2023). In our paper, we restrict attention to SA, since it is the predominant tax principle at the international level.

A number of recent studies empirically analyze the tax revenue (or welfare) consequences of the GMT in the context of fixed corporate tax rates. UNCTAD (2022) shows that the GMT (with the SBIE) could lead to a growth in global tax revenues generated by FDI income between 15 per cent and 20 per cent, if all host countries apply the QDMTT. Baraké et al. (2022) document that G7 countries could collect around €90 billion in the headquarters scenario, while it would fall to €17 billion under host country collection. By contrast, developing countries would favor the QDMTT over the IIR. Ferrari et al. (2023) model and quantify the effects of the GMT, showing that it can improve welfare in most countries by inducing higher tax revenues. Hugger et al. (2024) find that the GMT would narrow the tax differential and raise global tax revenues by between \$155 and \$192 billion on average per year. Very few empirical works take into account the tax rate adjustments by countries in response to the minimum tax. IMF (2023) estimate that the average tax rate would rise from 22.2% to 24.3% in response to the GMT, which in turn could increase global tax revenues by 8.1%. Buettner and Poehnlein (2024) examine the effects of a minimum tax on the tax policy of German municipalities in a context of local tax competition. After introducing the minimum tax, only high-tax municipalities reduce their business tax rates. The minimum tax does not affect high-tax municipalities' revenues but harms the tax havens.

We start from the short-run analysis, in which only the MNE can adjust its behavior. Introducing the GMT does not change the investment level in the large country. With the top-up tax paid by the low-tax affiliate, the GMT reduces the difference between tax rates on two affiliates' GloBE incomes (referred to as the *true* tax rate differential). Consequently, profit shifting is reduced, and the large country's tax revenue definitely increases. However, the revenue effect of the GMT is generally ambiguous for the small country. Starting from the equilibrium without the GMT, a marginal increase in the minimum rate has two opposite effects. Firstly, due to the SBIE, a higher GMT rate exerts investment incentives for the low-tax affiliate and thus raises the small country's revenue. Secondly, a higher GMT rate increases the top-up tax rate and incurs a larger revenue loss from the deduction of SBIE. Which effect can dominate depends on the initial tax rate of the small country (or equivalently, on the profit shifting cost of the MNE).

Then we investigate the long-run situation where both the MNE and the governments react to the tax reform. Countries compete in corporate tax rates to maximize their revenues while taking as given the international tax architecture. At the equilibrium of the tax game, the GMT does not necessarily bind the small country. It is binding only if the minimum tax is low. Otherwise, the small country will set its tax rate below the minimum, even at zero (if the carve-out rate is very small). The key insight is that due to tax deduction of the SBIE, lowering one country's tax rate below the GMT rate can incentivize real investment but also incurs a larger revenue loss since the top-up tax rate increases. For a high GMT rate, the former effect can dominate. Moreover, when the GMT rate is sufficiently high and the carve-out is not very small, both countries undercut the minimum so that profit shifting ends. In this case, countries aim at attracting capital investments instead of competing for paper profits.

Our results can be related to the traditional minimum taxation literature, which treats the minimum tax as a lower bound imposed on countries' tax rates. Kanbur and Keen (1993) initiatively explore the minimum tax in a commodity tax competition model with cross-border shopping. They show that both large and small countries can benefit from the minimum tax. More recently, Hebous and Keen (2023) extend this framework to study international taxation of MNEs and derive the levels of maximal Pareto dominant minimum tax rate and Pareto-efficient minimum rate. In the two papers, the minimum tax binds the small country and induces the large country to set tax rate along the unconstrained best response curve. Wang (1999) extends Kanbur and Keen's (1993) model to the Stackelberg tax-setting, and presents that the minimum tax not only binds the follower (i.e., the small country) but also may bind the leader (i.e., the large country). Besides, imposing a minimum tax constraint definitely benefits the leader and harms the follower. In our paper, we restrict attention to simultaneous move of countries, as is widely accepted in the profit shifting literature. The binding minimum rate is also an implicit assumption in Janeba and Schjelderup (2023), who study the GMT in a model where two identical non-haven countries compete for firms via tax rates or subsidies, with profits shifted to the tax haven. They

assume that the haven's tax rate adjusts to the minimum rate once the GMT is introduced, and show that the revenue effects of the GMT are ambiguous, depending on the fiscal instrument governments use. However, specifying the minimum tax as a constraint that no tax rate may be set below a minimum level somewhat deviates from the design of the GMT based on the OECD's Model Rules. The GMT under Pillar Two allows countries to collect additional revenue via top-up taxes when an affiliate's ETR falls below the minimum rate. Remarkably, Irish government has decided to keep its corporate tax rate at 12.5% and top up the rate to 15% for Irish affiliates of MNEs. It also states that Pillar Two will provide a sound and stable basis for inward investment into Ireland in the long-term (see Department of Finance, 2023). This is consistent with the argument in our paper that countries may undercut the GMT rate to promote real investments.

A few recent theoretical papers on the GMT go beyond “the minimum tax constraint assumption” and consider the case where countries' tax rates are below the GMT rate. Johannesen (2022) sets up a model of tax competition for paper profits among tax havens and non-haven countries, and assumes that the top-up taxes are collected by home countries (under the IIR). The GMT causes a loss of private consumption for the owners of the multinationals in non-haven countries, but also curbs profit shifting and boosts tax revenue. The net welfare effect is generally ambiguous for non-haven countries. Haufler and Kato (2024) develop a tax competition model, where a non-haven and a haven country are bound by the GMT rate for large MNEs, but can choose tax rates freely for small MNEs. They show that introducing a moderate minimum tax can raise tax revenues for both countries. As the GMT rate increases, each country has incentives to split the tax rate and set tax rate below the minimum for small MNEs. However, the SBIE is omitted in the two papers. Schjelderup and Stähler (2023) consider the SBIE in a standard MNE model where the host country's tax rate is below the GMT rate. They show that the SBIE works like a wage subsidy and investment subsidy for the low-tax affiliate, since it allows the firm to deduct payroll costs and user costs of tangible assets twice from the overall tax base. However, in their model countries' tax rates are assumed to be exogenously given, and the analysis of the revenue effect of the GMT is absent.

After characterizing the Nash equilibrium of the tax competition game, we examine the long-run revenue effects of the GMT. As in the short run, the GMT reduces the true tax rate differential between two countries and always benefits the large country. In contrast, the revenue effect for the small country is more subtle. We show that even a marginal GMT reform (with the minimum rate marginally above the small country's initial tax rate) may harm it. This may happen when the asymmetry between two countries is small and the profit shifting cost is intermediate. The key insight is that in this case the marginal reform induces both countries to undercut the minimum so that the profit shifted to the small country jumps *discontinuously* to zero. Since the small country attracts considerable paper profits absent the GMT, the revenue loss from eliminating profit shifting can outweigh the increased taxation on true profits generated by real economic activities. Furthermore, we provide conditions for

the Pareto-improving non-marginal tax reform. When the carve-out rate is not too small, the GMT rate is not very high and profit shifting is not very sensitive to the minimum, the GMT can benefit the small country by increasing the taxation of both true profit and shifted profit.

Our paper contributes to the emerging theoretical works on the GMT. First, we take into account the real investment responses of the MNE. In most theoretical GMT literature, firms' profits are assumed to be fixed and independent of tax rates so that only profit shifting behavior is considered. This assumption implies the adoption of a pure profit tax with full tax deductibility of costs, which is at odds with most countries' corporate tax systems. In contrast, our model allows for the partial deductibility of capital costs such that countries' tax policies affect the MNE's decision on both profit shifting and real investment. As is discussed below, partial deductibility creates scope for countries to gain tax revenue by incentivizing affiliates to increase investments. In this regard, our paper can be related to Chen and Hindriks (2023), who analyze the effects of tax deductibility in a model of tax competition with two countries and a tax haven. They derive conditions under which pure profit tax is superior (inferior) to turnover tax and investigate the optimal deductibility rate in the profit shifting context. In our paper, the (partial) deductibility rate is fixed and our focus is on the effects of the GMT rate and carve-out rate. Second, we capture the SBIE and QDMTT in the tax competition game. The SBIE is a key feature of Pillar Two but is largely ignored in the existing theoretical literature. Only a few papers theoretically investigate the effects of the SBIE (see Devereux et al., 2021, 2022; Schjelderup and Stähler, 2023), whereas the strategic interactions between countries' tax rates are overlooked in these works. Due to the SBIE, either increasing the minimum rate in the short run or setting tax rate below the minimum in the long run can exert investment incentives, while doing so also increases the top-up tax rate and incurs a larger revenue loss (see Sections 3.3 and 3.4 for details). This is the key trade-off in our analysis, which determines the short-run revenue effect of the GMT and reshapes the tax competition game in the long run. Third, our results are comparable with traditional minimum taxation literature. When all costs are tax deductible (which is equivalent to the assumption of fixed profits), investments in each country are undistorted at the initial equilibrium. The aforementioned tax policies fail to raise the affiliate's GloBE income through investment incentives, but only causes a larger revenue loss. Consequently, the GMT always harms the small country in the short run. In the long run, undercutting the minimum is a strictly dominated strategy for each country. Therefore, the GMT under Pillar Two works in the same way as "a constraint that no tax rate may be set below the minimum level". In this sense, the traditional minimum taxation model can be regarded as a special case in our paper.

The rest of the paper proceeds as follows. In Section 3.2, we present the model and analyze the equilibrium before the GMT is introduced. Section 3.3 analyzes the short-run effect of introducing the GMT. In Section 3.4, we characterize the Nash equilibrium of the tax competition game and investigate the long-run effects of the GMT. Section 3.5 concludes.

3.2. The model

Consider two asymmetric countries, labelled by 1 and 2, that form a small part of the world. Each country hosts an affiliate of a representative multinational enterprise (MNE). Each affiliate produces a homogenous good according to a decreasing returns to scale technology $f_i(k_i)$, where k_i is the capital employed by affiliate i . Decreasing returns to scale in production imply the existence of a fixed factor (e.g., entrepreneurial services) that generates economic rents. In Appendix 3.E, we briefly present the extended model that includes labor. Output is sold at the world market at a price normalized to unity. The two affiliates pay the (exogenously given) world interest rate r for per unit of capital use. For most corporate tax systems, the capital cost may not be fully tax deductible, since countries only allow the MNE to deduct the cost financed by debt but not by equity. Denote by $\mu \in [0,1)$ the fraction of capital cost that can be deducted from the corporate tax base. Although we mainly focus on the case of partial deductibility, our results also hold true for $\mu = 1$. As we will present in Sections 3.3 and 3.4, a pure profit tax with full deductibility – which is the underlying assumption in the existing minimum tax literature – can be regarded as a special case of our model. Since the main purpose of the paper is to investigate the effects of GMT, we impose a few structure on the production technology by specifying a quadratic production function: $f_i(k_i) := \alpha_i k_i - \frac{k_i^2}{2}$ with $\alpha_1 > \alpha_2 > r$.¹ In this setting, country 1 (country 2) is the large country (small country) in the sense that country 1 has a larger market size and a less elastic tax base.

The MNE can shift profits between two affiliates in order to minimize its tax liability. We abstract from the specific channels through which the MNE reallocates profits and denote the profit shifting level by g . If $g > 0$ ($g < 0$), then the MNE shifts profit to (from) country 2 and so the tax base in country 2 goes up (down). Profit shifting is costly for the MNE and involves a non-deductible concealment cost with the form $h(g) := \frac{\delta}{2} g^2$, $\delta > 0$. The concealment cost approach is widely used in the profit shifting literature (e.g., Devereux et al., 2008; Mardan and Stimmelmayer, 2018, 2020; Koethenbuerger et al., 2019; Janeba and Schjelderup, 2023). The tax rate of country i ($i = 1, 2$) is denoted by t_i . The GloBE income (i.e., taxable profit) of affiliate i is $\pi_i = f_i(k_i) - \mu r k_i + (-1)^i g$.

¹ The quadratic specification of a production function is usual in the profit shifting literature (e.g., Devereux et al., 2008; Johannesen, 2010; Haufler et al., 2018; Mardan and Stimmelmayer, 2020).

We consider a two-stage tax competition game. 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 real investment and profit shifting level to maximize total after-tax profits. The assumption of revenue-maximizing governments is frequently used in the international taxation literature (e.g., Johannesen, 2010; Mardan and Stimmelmayer, 2018; Koethenbueger et al., 2019; Janeba and Schjelderup, 2023; Haufler and Kato, 2024). It reflects the desire to increase tax payments from MNEs and is line with the objective of the OECD/G20 Inclusive Framework on BEPS. Tax revenue considerations play a very important role in the taxation of MNEs, due to the low taxes that many large MNEs are paying. In reality, profit shifting causes severe revenue shortfalls and raises equality-of-treatment concerns in many countries, which motivates politicians to stabilize the revenue from corporate taxes.

Without the GMT, the MNE's after-tax profit is:

$$\Pi = \sum_{i=1}^2 \left[(1-t_i) \left(f_i - \mu r k_i + (-1)^i g \right) - (1-\mu) r k_i \right] - \frac{\delta}{2} g^2.$$

Solving the MNE's profit maximization leads to:

$$k_i = \max \left\{ \frac{\alpha_i (1-t_i) - (1-\mu t_i) r}{1-t_i}, 0 \right\}, \quad i = 1, 2, \quad (3.1)$$

$$g = (-1)^i \min \left\{ \frac{t_j - t_i}{\delta}, f(k_j) - \mu r k_j \right\} \quad \text{if } t_j \geq t_i. \quad (3.2)$$

From (3.1), for interior solution with $k_i > 0$, $\frac{\partial k_i}{\partial t_i} < 0$. Intuitively, a higher tax rate

increases the tax burden of the affiliate and so discourages real investment. From (3.2), the MNE shifts profit from high-tax country to low-tax country. In particular, the high-tax affiliate reports zero taxable profit if the tax differential is sufficiently large. (3.1) and (3.2) replicate the standard results in the profit shifting literature.

The tax revenue function of country i reads:

$$R_i(t_i, t_j) = t_i \pi_i = t_i \left(f_i - \mu r k_i + (-1)^i g \right). \quad (3.3)$$

Due to the complexity inherent in partial tax deductibility and asymmetry settings, the model has no closed-form solution for all $\mu \in [0, 1)$. In what follows, we assume that

$$\frac{\alpha_1 - r}{\alpha_1 - \mu r} \leq 2 \frac{\alpha_2 - r}{\alpha_2 - \mu r} \quad (\text{or equivalently, } \alpha_2 := \frac{r[\alpha_1(2 - \mu) - \mu r]}{\alpha_1 + r - 2\mu r} \leq \alpha_2 < \alpha_1). \quad \text{It means that}$$

country 2 is not “too small” relative to country 1. As shown in Appendix 3.A.1, under this assumption country 2 has no incentive to become a tax haven with shifted profits being its only tax base. Denote by t_i^N country i ’s equilibrium tax rate absent the GMT. Then we can present the following:

Lemma 1. *Before the introduction of the GMT,*

- (i) *there exists a unique Nash equilibrium with $t_i^N \in \left(0, \frac{\alpha_i - r}{\alpha_i - \mu r}\right)$, $i = 1, 2$;*
- (ii) *at equilibrium, the small country always undercuts the large country, i.e., $t_1^N > t_2^N$.*

Proof. See Appendices 3.A.1 and 3.A.3.

The lemma establishes the existence and uniqueness of the Nash equilibrium absent the GMT. There are capital investments in both countries at equilibrium, while the small country sets a lower tax rate to attract profits from the large country. Intuitively, under equal tax rates $t_1 = t_2$ the small country has a higher tax base elasticity (in absolute value) than the large country (i.e., $-\frac{\partial \pi_2}{\partial t_2} \frac{t_2}{\pi_2} > -\frac{\partial \pi_1}{\partial t_1} \frac{t_1}{\pi_1}$). Hence, the small country’s tax base is more sensitive to tax rate changes, which forces it to tax less.

Remark 1. Using (A4) and Lemma 1(ii) leads to $\frac{\partial t_1^N}{\partial \delta} > 0$, i.e., the equilibrium tax rate of country 1 strictly increases with the profit shifting cost. Denote by $\bar{t}_1 := \lim_{\delta \rightarrow +\infty} t_1^N(\delta)$ the (least) upper bound for t_1^N . In contrast, the effect of profit shifting cost on country 2’s equilibrium tax rate is ambiguous. Numerical simulations indicate that t_2^N may decrease with δ when δ is large.

Remark 2. When $\delta \rightarrow +\infty$, the profit shifting is eliminated so that the positive tax externality vanishes. In this limiting case, by setting the efficient tax rate $\bar{t}_1$, country 1 can achieve the first-best solution $\bar{R}_1 := \lim_{\delta \rightarrow +\infty} R_1^N(t_1^N(\delta), t_2^N(\delta))$ absent the GMT. Formally, we have the following:

$$R_1(t_1^N, t_2^N) = \varphi_1(t_1^N) - \frac{t_1^N(t_1^N - t_2^N)}{\delta} < \varphi_1(t_1^N) < \varphi_1(\bar{t}_1) = \bar{R}_1, \quad (3.4)$$

where $\varphi_1(t_1^N) := t_1^N(f_1(k_1^N) - \mu r k_1^N)$ denotes the revenue from taxing affiliate 1's true profit generated by substantive activities.

(3.4) indicates that $\bar{R}_1$ is the upper bound for country 1's equilibrium revenue without the GMT. Nevertheless, after introducing the GMT, country 1 can receive an equilibrium revenue larger than $\bar{R}_1$ when the minimum is sufficiently high and the carve-out is not very small (see Section 3.4 and Appendix 3.A.7.4).

3.3. Short-run analysis of the GMT: fixed tax rates

After introducing the GMT, affiliate i 's GloBE income π_i is targeted for additional taxation (i.e., top-up tax) when country i 's tax rate falls below the GMT rate t_m . As in the previous minimum taxation literature (e.g., Kanbur and Keen, 1993; Wang, 1999; Keen and Konrad, 2013), we assume throughout the paper that the GMT rate lies between the initial tax rates of two countries, i.e., $t_m \in (t_2^N, t_1^N)$. The top-up tax rate for affiliate i is $\max\{t_m - t_i, 0\}$, which is the difference between the minimum rate and the host country's tax rate. The SBIE allows the MNE to tax-deduct a fraction of the capital stock from the GloBE income, which reduces the top-up tax liability of the affiliate with substantive activity. The GloBE income after the deduction of the SBIE is the excess profit $E_i = \pi_i - \sigma k_i$, where σ denotes the carve-out rate and σk_i is the SBIE. The excess profit constitutes the tax base for

the top-up tax. So the top-up tax owed by affiliate i is $\max\{t_m - t_i, 0\} \cdot (\pi_i - \sigma k_i)$. As mentioned in Section 3.1, an MNE's tax liability is the same, whether the top-up tax is collected by the headquarters country (under the IIR) or by the host country (under the QDMTT). The low-tax country has very strong incentives to implement the QDMTT, because otherwise it would leave "money on the table" for other countries without changing the MNE's tax burden (see also Perry, 2023). In light of this, we assume throughout the paper that each country adopts the QDMTT to collect top-up taxes from the affiliate that is recording undertaxed profits in its territory.

In the short run, two countries' tax rates remain unchanged, while the MNE is able to adjust the decisions on investment and profit shifting. This reflects the fact that governments usually need some time to adjust tax rates in response to the tax reform.

The after-tax profit of the MNE reads:

$$\begin{aligned} \Pi^m &= \sum_{i=1}^2 \left[(1-t_i^N) \underbrace{(f_i - \mu r k_i + (-1)^i g)}_{\text{GloBE income}} - (1-\mu) r k_i \right] - \underbrace{(t_m - t_2^N)(f_2 - \mu r k_2 + g - \sigma k_2)}_{\text{top-up tax}} - \frac{\delta}{2} g^2 \\ &= (1-t_1^N)(f_1 - \mu r k_1 - g) - (1-\mu) r k_1 + (1-t_m)(f_2 - \mu r k_2 + g) - (1-\mu) r k_2 \\ &\quad + \underbrace{(t_m - t_2^N)\sigma k_2}_{\text{tax saved from the SBIE}} - \frac{\delta}{2} g^2 \end{aligned} \quad , \quad (3.5)$$

where superscripts m represents the introduction of the GMT.

As shown by (3.5), with the additional GloBE top-up tax, affiliate 2's GloBE income is now taxed at the minimum rate t_m . On the other hand, the SBIE reduces the exposure to the minimum tax so that the tax amount of $(t_m - t_2^N)\sigma k_2$ can be saved by affiliate 2.

The first-order conditions of the profit maximization are:

$$\frac{\partial \Pi^m}{\partial k_1} = (1-t_1^N)(f_1' - \mu r) - (1-\mu)r = 0 \Rightarrow k_1^m = \frac{\alpha_1(1-t_1^N) - (1-\mu t_1^N)r}{1-t_1^N} = k_1^N, \quad (3.6)$$

$$\frac{\partial \Pi^m}{\partial k_2} = (1-t_m)(f_2' - \mu r) - (1-\mu)r + \sigma(t_m - t_2^N) \leq 0 \Rightarrow$$

$$k_2^m = \max \left\{ \frac{\alpha_2(1-t_m) - (1-\mu t_m)r + (t_m - t_2^N)\sigma}{1-t_m}, 0 \right\}, \quad (3.7)$$

$$\frac{\partial \Pi^m}{\partial g} = t_1^N - t_m - \delta g = 0 \quad \Rightarrow \quad g^m = \frac{t_1^N - t_m}{\delta}. \quad (3.8)$$

Notably, for a high carve-out rate, affiliate 2's excess profit E_2 will be negative so that the GMT is immaterial. Since the main focus of this paper is the effects of the GMT, we restrict attention to positive excess profit and assume that the short-run carve-out rate satisfies

$$\sigma \leq \frac{\alpha_2(1-t_m) + r(1-(2-t_m)\mu)}{2-t_2^N-t_m} =: \bar{\sigma}^S(t_m), \text{ where superscript } S \text{ represents the short-run}$$

analysis. As shown in Appendix 3.B.1, this assumption ensures that $E_2 > 0$.

The GMT is inactive for affiliate 1, since country 1's tax rate is above the minimum. As shown by (3.6), the capital investment in country 1 is unaffected. In contrast, increasing the GMT rate has two opposite effects on the capital investment of affiliate 2. Specifically, for interior solution with $k_2^m > 0$, differentiating (3.7) with respect to t_m yields:

$$\frac{\partial k_2^m}{\partial t_m} = \underbrace{\frac{\partial}{\partial t_m} \left[\frac{\alpha_2(1-t_m) - (1-\mu t_m)r}{1-t_m} \right]}_{\text{tax burden effect}} + \underbrace{\frac{\partial}{\partial t_m} \left[\frac{(t_m - t_2^N)\sigma}{1-t_m} \right]}_{\text{tax incentive effect}} = \frac{\sigma(1-t_2^N) - r(1-\mu)}{(1-t_m)^2}.$$

The first term on the right-hand side is negative. A higher GMT rate increases the taxation on affiliate 2's GloBE income and tends to reduce the real investment in country 2. We refer to it as the "tax burden effect". The second term is positive and captures the "tax incentive effect". A higher GMT rate means that more taxes can be saved due to the deduction of the SBIE (cf. (3.5)), which incentivizes affiliate 2 to employ more capital. Moreover, a higher carve-out rate makes the tax incentive effect stronger, since it leads to a

larger SBIE. When $\sigma > \frac{r(1-\mu)}{1-t_2^N}$ ($\sigma < \frac{r(1-\mu)}{1-t_2^N}$), the second (first) effect dominates such

that country 2's investment level increases (decreases) with the minimum rate.

(3.8) indicates that the GMT reduces the profit shifting of the MNE. From (3.5), the MNE's profit shifting decision depends on the difference between the tax rates on two affiliates' GloBE incomes (taking into account the top-up tax paid by the low-tax affiliate). Henceforth, we refer to it as the *true* tax rate differential between two countries. After introducing the GMT, the tax rate on affiliate 2's GloBE income increases from t_2^N to t_m .

So the true tax differential narrows down and profit shifting decreases.

The short-run tax revenues of two countries are:

$$R_1^m = t_1^N (f_1 - \mu r k_1 - g), \quad (3.9)$$

$$\begin{aligned} R_2^m &= t_2^N (f_2 - \mu r k_2 + g) + (t_m - t_2^N)(f_2 - \mu r k_2 + g - \sigma k_2) \\ &= t_m (f_2 - \mu r k_2 + g) - \underbrace{(t_m - t_2^N) \sigma k_2}_{\text{loss from the deduction of SBIE}}. \end{aligned} \quad (3.10)$$

As shown by (3.10), the introduction of the GMT does not bring country 2's total taxes to the minimum rate t_m . The tax amount (i.e., $(t_m - t_2^N) \sigma k_2$ in (3.5)) saved by affiliate 2 due to the SBIE corresponds to a revenue loss for country 2.

The following proposition presents the revenue effect of the GMT when countries' tax rates are fixed.

Proposition 1. *In the short run where countries' tax rates are fixed,*

- (i) *the large country benefits from the GMT;*
- (ii) *introducing a GMT with minimum rate marginally higher than the small country's equilibrium tax without the GMT increases (reduces) the small country's tax revenue if $t_2^N > t_2^*$ ($t_2^N < t_2^*$);*
- (iii) *(the non-marginal reform) assuming that the small country's revenue function is quasiconcave in t_m , then the GMT always reduces its tax revenue if $t_2^N < t_2^*$,*

$$\text{where } t_2^* := 1 - \sqrt{\frac{r(1-\mu)}{\alpha_2 - \mu r}}.$$

Proof. See Appendix 3.A.4.

Intuitively, the introduction of GMT does not change the investment level in country 1 but reduces the outward profit shifting. So the GMT raises the large country's tax revenue. In contrast, the revenue effect of the GMT is ambiguous for the small country. Increasing the minimum rate marginally above the initial tax rate of country 2 affects its tax revenue in the following way:

$$\left. \frac{\partial R_2^m}{\partial t_m} \right|_{t_m=t_2^N} = \underbrace{t_2^N (f_2' - \mu r) \sigma}_{\text{gain from tax incentive}} - \underbrace{\sigma k_2^N}_{\text{loss from SBIE}}.$$

The first term is positive and captures the revenue gain from investment incentives. Recall that due to the SBIE, a higher GMT rate has a tax incentive effect, which exerts investment incentives for affiliate 2, increases its GloBE income and thus raises country 2's revenue. The second term is negative. All else equal, a higher GMT rate increases the top-up tax rate and incurs a larger revenue loss for country 2 due to the deduction of SBIE (cf. (3.10)). Moreover, for a higher initial tax rate t_2^N , the tax incentive effect arising from the marginal rise of the minimum is stronger, while country 2's capital stock and the revenue loss from the SBIE are smaller. Therefore, the revenue gain generated by the marginal tax reform can dominate if and only if country 2's initial tax rate is high.

In the special case of pure profit tax ($\mu=1$), the threshold $t_2^*=1$. Then by the proposition, the marginal tax reform always harms the small country. For a pure profit tax, all costs are deductible and investment in each country is undistorted at the initial equilibrium, with the marginal product of capital equaling the exogenous interest rate. Consequently, a higher GMT rate fails to raise affiliate 2's GloBE income through the investment incentives. The marginal reform only causes a larger revenue loss from the deduction of SBIE and thus reduces the small country's revenue.

Part (iii) further shows the global property of the GMT for any minimum rate on the interval (t_2^N, t_1^N) , and is more relevant to practical policies. It states that the GMT generally reduces the short-run revenue of the small country if its initial tax rate is low. In Appendix 3.C, we provide sufficient conditions for the quasiconcavity of $R_2^m(t_m)$. Specifically, if

$$\sigma \leq \frac{r(1-\mu)}{1-t_2^N} \quad \text{or if} \quad \frac{r(1-\mu)(2+t_2^N)}{(1-t_2^N)(2-t_2^N)} \leq \sigma \leq \bar{\sigma}^S(t_1^N), \quad \text{then } R_2^m(t_m) \text{ is quasiconcave in } t_m \text{ for all } t_m \in (t_2^N, t_1^N).$$

Lastly, the following lemma reveals how the profit shifting cost affects the comparison of t_2^N and t_2^* .

Lemma 2. *There exists a threshold of the concealment cost parameter δ^* such that $t_2^N > t_2^*$ ($t_2^N \leq t_2^*$) if $\delta > \delta^*$ ($\delta \leq \delta^*$).*

Proof. See Appendix 3.A.5.

Unlike country 1, country 2's equilibrium tax rate does not necessarily monotonically increase with the concealment cost (see Remark 1). However, t_2^* must intersect $t_2^N(\delta)$ at the upward-sloping part of $t_2^N(\delta)$ (see Appendix 3.A.5). This property directly leads to the lemma.

It immediately follows from Proposition 1 and Lemma 2 that a marginal reform of the GMT will raise (reduce) the small country's short-run revenue under high (low) profit shifting cost. Moreover, assuming the quasiconcavity of the small country's revenue function, then introducing a GMT with minimum rate lying between two countries' initial tax rates always harms the small country in low concealment cost environments.

3.4. Long-run analysis: tax competition

In the long run, both the MNE and the governments adjust their strategies in response to the GMT. In the first stage, two countries simultaneously choose their tax rates, while taking the minimum and carve-out rate as given and anticipating the MNE's responses to their tax choices. In the second stage, the MNE chooses capital investment in each country and the profit shifting level. In the long-run analysis, we assume that the carve-out rate $\sigma \in (\underline{\sigma}, \bar{\sigma}]$,

where $\underline{\sigma} := \frac{(1 - \mu t_m)r - \alpha_2(1 - t_m)}{t_m}$ and $\bar{\sigma} := \frac{\alpha_2(1 - t_m) + r[1 + (t_m - 2)\mu]}{2 - t_m}$. As is shown below,

the lower bound on the carve-out rate ensures that the small country can attract some investment by setting a low tax rate. It rules out the case where country 2 becomes a tax haven with $k_2 \equiv 0$, $\forall t_2 \in [0, 1]$. In Appendix 3.D, we analyze $\sigma \leq \underline{\sigma}$ (which corresponds to a high minimum together with an overly small carve-out) and show that there exists a continuum of Nash equilibria in this case. The upper bound on the carve-out ensures that the excess profit is positive when each country sets its tax below the GMT rate, i.e., $E_i := \pi_i - \sigma k_i \geq 0$, $\forall t_i \in [0, t_m)$, $\forall i \in \{1, 2\}$ (see Appendix 3.B.2). It rules out the case where the minimum tax is inconsequential for country i due to negative excess profit.

3.4.1. The equilibrium analysis

We first investigate the situation where one country's tax rate is below the GMT rate. Suppose that country i 's tax rate $t_i < t_m$. Then the after-tax profit of the MNE reads:

$$\begin{aligned} \Pi^m = & (1-t_m) \left(f_i - \mu r k_i + (-1)^i g \right) - (1-\mu) r k_i + \underbrace{(t_m - t_i) \sigma k_i}_{\text{tax saved from the SBIE}} + (1-t_j) \left(f_j - \mu r k_j + (-1)^j g \right) \\ & - (1-\mu) r k_j - \mathbf{1}_j (t_m - t_j) \left(f_j - \mu r k_j + (-1)^j g - \sigma k_j \right) - \frac{\delta}{2} g^2 \end{aligned} \quad (3.11)$$

where indicator variable $\mathbf{1}_j$ equals unity if $t_j < t_m$ and zero otherwise.

Solving the profit maximization of the MNE yields:

$$k_i^m = \max \left\{ \frac{\alpha_i(1-t_m) - (1-\mu t_m)r + (t_m - t_i)\sigma}{1-t_m}, 0 \right\}, \quad (3.12)$$

$$k_j^m = \begin{cases} \max \left\{ \frac{\alpha_j(1-t_j) - (1-\mu t_j)r}{1-t_j}, 0 \right\} & \text{if } t_j \geq t_m \\ \max \left\{ \frac{\alpha_i(1-t_m) - (1-\mu t_m)r + (t_m - t_i)\sigma}{1-t_m}, 0 \right\} & \text{if } t_j < t_m \end{cases}, \quad (3.13)$$

$$g^m = \begin{cases} (-1)^i \min \left\{ \frac{t_j - t_m}{\delta}, f_j(k_j^m) - \mu r k_j^m \right\} & \text{if } t_j \geq t_m \\ 0 & \text{if } t_j < t_m \end{cases}. \quad (3.14)$$

From (3.12), for interior solution with $k_i^m > 0$, we have $\frac{\partial k_i^m}{\partial t_i} < 0$ and $\frac{\partial k_i^m}{\partial \sigma} > 0$. The intuition is straightforward. Due to the deduction of SBIE, affiliate i can reduce its tax payments by $(t_m - t_i)\sigma k_i$, as presented by (3.11). Hence, either a lower tax rate of country i or a higher carve-out rate increases the tax incentives and raises the capital affiliate i employs. In particular, by reducing its tax rate down to zero, country i can gain a maximum level of investment with $\bar{k}_i := \frac{\alpha_i(1-t_m) - (1-\mu t_m)r + t_m \sigma}{1-t_m}$. Under the assumption $\sigma > \underline{\sigma}$, we

have $\bar{k}_2 > 0$, which means that the small country is able to attract inward investment by setting a low corporate tax rate. Combining (3.13) and (3.14) indicates that for all $t_i < t_m$, the

profit shifting of the MNE is *independent* of country i 's tax choice. Whenever country i sets its tax rate below the minimum, the top-up tax is triggered so that affiliate i 's GloBE income is always taxed at t_m (cf. (3.11)). Country i cannot attract more paper profits from country j by further lowering its tax rate. So the GMT places a floor on the taxation of GloBE incomes. It reduces countries' incentives to compete for paper profits and mitigates the corporate tax competition. Moreover, when both countries' tax rates are below the GMT rate, two affiliates' GloBE incomes are taxed at the same rate of t_m , so that the true tax differential vanishes. In this case, profit shifting is irrelevant to the MNE's total tax liability and can be eliminated.

Under the QDMTT, Country i 's tax revenue function is:

$$\begin{aligned} R_i^m(t_i, t_j) &= t_i \left(f_i - \mu r k_i + (-1)^i g \right) + (t_m - t_i) \left(f_i - \mu r k_i + (-1)^i g - \sigma k_i \right) \\ &= t_m \left(f_i - \mu r k_i + (-1)^i g \right) - \underbrace{(t_m - t_i) \sigma k_i}_{\text{loss from the deduction of SBIE}}. \end{aligned} \quad (3.15)$$

The following lemma presents the effect of one country's tax rate when it is below the minimum rate.

Lemma 3. *Suppose that country i 's tax rate is below the GMT rate, i.e., $t_i < t_m$. Then country i 's tax rate affects its tax revenue in the following way:*

- (i) For $t_m \leq t_i^*$, country i 's revenue strictly increases with t_i for all $t_i \in [0, t_m)$;
- (ii) For $t_m > t_i^*$ and $\underline{\sigma} < \sigma \leq \sigma_i^m$, country i 's revenue decreases with t_i for all $t_i \in [0, t_m)$;
- (iii) For $t_m > t_i^*$ and $\max\{\underline{\sigma}, \sigma_i^m\} < \sigma \leq \bar{\sigma}$, country i 's revenue increases (decreases)

$$\text{with } t_i \text{ when } t_i \in \left[0, \left(1 - \frac{\sigma_i^m}{\sigma} \right) t_m \right) \quad (t_i \in \left(\left(1 - \frac{\sigma_i^m}{\sigma} \right) t_m, t_m \right)),$$

$$\text{where } t_i^* := 1 - \sqrt{\frac{r(1-\mu)}{\alpha_i - \mu r}} \text{ and } \sigma_i^m := \frac{r(1 - 2\mu t_m + \mu t_m^2) - \alpha_i(1 - t_m)^2}{t_m(2 - t_m)}.$$

Proof. See Appendix 3.A.6.

Part (i) of Lemma 3 implies that one country will never set its tax rate below the minimum when the minimum rate is low. Otherwise, as shown by part (ii) and (iii), it still has incentives to keep a tax rate below the minimum, even possibly at zero. The intuition is as follows. Given (3.15), starting from the minimum rate, a marginal decrease in one country's tax rate affects its revenue in the following way:

$$-\frac{\partial R_i^m}{\partial t_i} \Big|_{t_i=t_m} = \underbrace{-t_m (f_i' - \mu r) \frac{\partial k_i}{\partial t_i}}_{>0} - \sigma k_i. \quad (3.16)$$

The first effect in (3.16) is positive. From (3.12), a lower tax rate of country i incentivizes affiliate i to employ more capital. This increases affiliate i 's GloBE income and tends to raise country i 's tax revenue. The second effect in (3.16) is negative. Reducing country i 's tax rate increases the top-up tax rate and thus causes a larger revenue loss for country i due to the deduction of SBIE. Furthermore, a higher t_m renders real investment

more sensitive to the tax rate change (i.e., $\frac{\partial}{\partial t_m} \left| \frac{\partial k_i}{\partial t_i} \right| > 0$) and so makes the first effect stronger.

Therefore, when t_m is high, the first effect prevails so that one country can increase its tax revenue by undercutting the minimum. For a low t_m , the second effect dominates so that any tax choice below the minimum is a strictly dominated strategy.

We now analyze the role of carve-out rate. For $t_m > t_i^*$, assume that country i sets its tax rate below the minimum. Recall from (3.12) that either a higher carve-out rate or a lower tax rate promotes the investment. So when the carve-out rate decreases, country i has to reduce its tax rate to offset the negative effect on capital investment in order to maintain the revenue-maximizing capital level with $k_i^m = \frac{\alpha_i - r}{2 - t_m}$ (see Appendix 3.A.6). In particular, for a

very small carve-out rate with $\sigma \leq \sigma_i^m$, country i has to reduce its tax rate down to zero to maximize its revenue.

Notably, it directly follows from Lemma 3(ii) and (iii) that: for $t_m > t_i^*$ and $t_j \in [0,1]$,

$$\arg \max_{t_i \in [0, t_m]} R_i^m(t_i, t_j) = \max \left\{ 0, \left(1 - \frac{\sigma_i^m}{\sigma} \right) t_m \right\} =: \tilde{t}_i. \text{ Denote by } t_i^m \text{ the Nash equilibrium tax rate}$$

of country i after the GMT is introduced. We can establish the following:

Proposition 2. *After the introduction of the GMT, two countries set equilibrium tax rates in the following way:*

- (i) For $t_m \leq t_2^*$, $t_1^m = t_1(t_m)$, $t_2^m = t_m$;
- (ii) For $t_2^* < t_m \leq t_1^*$, $t_1^m = t_1(t_m)$, $t_2^m = \tilde{t}_2$;
- (iii) For $t_m > t_1^*$, two countries' equilibrium taxes are $(t_1^m = t_1(t_m), t_2^m = \tilde{t}_2)$ when $R_1(t_1(t_m), t_m) > R_1^m(\tilde{t}_1, \tilde{t}_2)$, and $(t_1^m = \tilde{t}_1, t_2^m = \tilde{t}_2)$ when $R_1(t_1(t_m), t_m) < R_1^m(\tilde{t}_1, \tilde{t}_2)$; both $(t_1(t_m), \tilde{t}_2)$ and $(\tilde{t}_1, \tilde{t}_2)$ are equilibrium tax rates when $R_1(t_1(t_m), t_m) = R_1^m(\tilde{t}_1, \tilde{t}_2)$.
- (iv) A sufficient condition for $(\tilde{t}_1, \tilde{t}_2)$ to be Nash equilibrium is: $t_m > t^{**}$ and $\max\{\underline{\sigma}, \sigma_1^m\} < \sigma \leq \bar{\sigma}$,

$$\text{where } R_1(t_1(t_m), t_m) = \varphi_1(t_1(t_m)) - \frac{t_1(t_m)[t_1(t_m) - t_m]}{\delta},$$

$$R_1^m(\tilde{t}_1, \tilde{t}_2) = \begin{cases} \frac{(\alpha_1 - r)^2}{2(2 - t_m)} & \text{if } \sigma > \sigma_1^m \\ \frac{(\alpha_1 - r)^2}{2(2 - t_m)} - \frac{(\sigma_1^m - \sigma)^2(2 - t_m)t_m^2}{2(1 - t_m)^2} & \text{if } \sigma \leq \sigma_1^m \end{cases}, \text{ and } t^{**} \text{ is determined by}$$

$$\frac{(\alpha_1 - r)^2}{2(2 - t^{**})} = \bar{R}_1 \text{ with } t^{**} \in (t_1^*, \bar{t}_1).$$

Proof. See Appendix 3.A.7.

Remark 3. When $t_m > t_1^*$ and $R_1(t_1(t_m), t_m) = R_1^m(\tilde{t}_1, \tilde{t}_2)$, the two Nash equilibria are Pareto-rankable. Specifically, $(t_1(t_m), \tilde{t}_2)$ Pareto-dominates $(\tilde{t}_1, \tilde{t}_2)$. The two equilibria involve the same investment level in country 2. However, country 2 has additional revenue at the former equilibrium by taxing the profit shifted from country 1, since the two affiliates' GloBE incomes are taxed at different rates. By contrast, the profit shifting is eliminated at the latter equilibrium. So we have: $R_2^m(\tilde{t}_2, t_1(t_m)) > R_2^m(\tilde{t}_2, \tilde{t}_1)$.

Remark 4. An immediate consequence of Proposition 2 is that the GMT reduces the MNE's profit shifting at equilibrium. When $t_m \leq t_2^*$, the GMT binds country 2 and the tax rate differential is $t_1(t_m) - t_m$. When $t_m > t_2^*$, country 2 undercuts the minimum and triggers the top-up tax so that affiliate 2's GloBE income is still taxed at t_m (cf. (3.11)). The true tax rate differential is $t_1(t_m) - t_m$ if country 1 sets the tax rate $t_1(t_m)$, and zero if country 1 also undercuts the minimum. Furthermore, $t_1(t_m) - t_m$ strictly decreases with the minimum rate, since $t_1'(t_m) < 1$ by (A1). So the GMT always narrows the true tax differential at equilibrium and curbs profit shifting.

Our results are substantially different from the traditional minimum taxation literature which treats minimum tax as a lower bound imposed on each country's tax rate. In such settings, the minimum tax always binds the low-tax country, inducing the high-tax country to set tax rate along the unconstrained best response curve (see Kanbur and Keen, 1993; Keen and Konrad, 2013; Hebous and Keen, 2023). The binding minimum rate is also an implicit assumption in Janeba and Schjelderup (2023), who model the GMT as an increase in the tax haven's tax rate. In our paper, for a pure profit tax ($\mu = 1$), the threshold $t_i^* = 1$. Then by Proposition 2(i), the GMT is binding for the small country. For each country, the capital investment is undistorted whenever its pure profit tax rate is above the minimum. Then lowering the tax rate marginally below the minimum is unable to alter the affiliate's GloBE income through investment incentives, but only causes a larger revenue loss (cf. (3.16)). So neither country has incentives to undercut the minimum. In this special case, the GMT under Pillar Two works in the same way as “a constraint that no tax rate may be set below the minimum level”.

However, for the general case of partial deductibility ($\mu < 1$), the GMT does not necessarily bind the small country. It is binding only if the minimum rate is low. Otherwise, the small country will set its tax rate below the minimum. The intuition is the following. Starting from a high minimum rate with $t_m > t_2^*$, country 2 can be better off by undercutting the minimum, since the revenue gain from larger real investments outweighs the revenue loss from the deduction of SBIE (see the discussion below Lemma 3). On the other hand, recall that absent the GMT, country 2's tax base is very sensitive to tax rate changes due to the small market size. Consequently, given country 1's equilibrium tax rate, the tax revenue curve of country 2 is downward-sloping for all $t_2 \geq t_m$. Starting from the minimum rate, choosing a higher tax rate always reduces country 2's revenue. So when the minimum rate is above the threshold t_2^* , country 2's best response is to undercut the minimum.

In contrast, the GMT shapes the large country's revenue function in a different way. Since country 1 has a larger market size, its tax base is less sensitive to tax rate changes absent the GMT. Therefore, given country 2's equilibrium tax, country 1's revenue curve is upward-sloping for all $t_1 \in [t_m, t_1(t_m)]$. Starting from the minimum rate, it can be better off by choosing a higher rate $t_1(t_m)$. On the other hand, for a high minimum rate with $t_m > t_1^*$, undercutting the minimum is also beneficial for country 1 due to larger inward investments. This sheds new light on the “conventional wisdom” that countries' revenue functions (i.e., the Laffer curves) are unimodal and quasiconcave. In our paper, when $t_m > t_1^*$, country 1's revenue function $R_1^m(t_1, t_2^m)$ becomes *bimodal* with two peaks and is *non-quasiconcave* in t_1 . In this case, country 1 chooses between undercutting the minimum (with the GloBE income taxed at the minimum rate and more inward investments) and setting a high tax rate $t_1(t_m)$ on the GloBE income (at the cost of smaller investment and outward profit shifting), depending on which tax revenue is larger. In particular, when the GMT rate is sufficiently high and the carve-out is not very small, the benefit of attracting larger investments dominates that of levying a higher tax rate so that country 1 will also undercut the minimum and collect top-up taxes (see Proposition 2(iv)).

Notably, one country's equilibrium tax rate is *decreasing* in the GMT rate, whenever it is below the minimum, i.e., $\frac{\partial}{\partial t_m} \left[\left(1 - \frac{\sigma_i^m}{\sigma} \right) t_m \right] < 0$. Recall that a higher GMT rate makes capital

investment more sensitive to tax rate changes (i.e., $\frac{\partial}{\partial t_m} \left| \frac{\partial k_i}{\partial t_i} \right| > 0$). This increases the benefit

of undercutting the minimum and incentivizes the country to reduce its tax rate, even down to zero (if the carve-out rate is small). While country 2's tax rate is continuous in the minimum, gradually increasing the GMT rate may trigger country 1's tax rate to jump *discontinuously* below the minimum rate. In this case, both countries aim at attracting real investment instead of competing for paper profits. In summary, the GMT under Pillar Two will reshape the underlying tax game and the competition pattern between two countries.

3.4.2. The revenue effects of the GMT

In this subsection, we investigate how the GMT affects countries' tax revenues in the long run. First, we show how the comparison of t_2^N and t_1^* is related to the market-size asymmetry and profit shifting cost.

Lemma 4. (i) For $\alpha_2 > \alpha_2^*$, there exists a threshold of the concealment cost parameter δ^{**} with $\delta^{**} > \delta^*$ such that $t_2^N > t_1^*$ ($t_2^N \leq t_1^*$) if $\delta > \delta^{**}$ ($\delta \leq \delta^{**}$); (ii) For $\alpha_2 \leq \alpha_2^*$, $t_2^N < t_1^*$ always holds, where $\alpha_2^* := \sqrt{(\alpha_1 - \mu r) \left[2\sqrt{r(1-\mu)(\alpha_1 - \mu r)} - r(1-\mu) \right]} + \mu r \in (\underline{\alpha}_2, \alpha_1)$.

Proof. See Appendix 3.A.8.

Absent the GMT, a reduction in country 2's market size increases the elasticity of its tax base (in absolute value) and thus imposes a downward pressure on its tax rate. When country 2 is sufficiently small (i.e., the market-size asymmetry is sufficiently large), this pressure is so high that country 2's equilibrium tax rate is always below t_1^* . Otherwise, t_1^* intersects $t_2^N(\delta)$ at the upward-sloping part of $t_2^N(\delta)$, which directly leads to part (i) of the lemma.

The following proposition reveals how the GMT affects equilibrium tax revenues after the behavioral adjustments of the MNE and two countries in response to the tax reform.

Proposition 3. *In the long run where countries' tax rates are endogenously determined,*

- (i) *the GMT always increases the large country's tax revenue.*
- (ii) *introducing a GMT with minimum rate marginally higher than the small country's equilibrium tax without the GMT increases the small country's tax revenue when $\alpha_2 \in [\underline{\alpha}_2, \alpha_2^*]$, or when $\alpha_2 \in (\alpha_2^*, \alpha_1)$ and $\delta \leq \delta^{**}$. Otherwise, it may reduce the small country's tax revenue.*
- (iii) *(the non-marginal reform) the GMT can raise the small country's tax revenue whenever the carve-out and minimum rate satisfy $\sigma \in [\sigma_2^m, \bar{\sigma}]$, $t_m \leq t_1^*$ and $\varepsilon_g(t_m) \in (0, 1]$, where $\varepsilon_g(t_m) := -\frac{\partial g^m / g^m}{\partial t_m / t_m}$ is the elasticity of profit shifting with respect to the GMT rate (in absolute value).*

Proof. See Appendix 3.A.9.

Remark 5. If the small country loses from a marginal GMT reform in the short run, it can be better off in the long run. This result directly follows from Proposition 1(ii), Lemma 2 and Proposition 3(ii). Introducing the marginal reform in an international tax environment with low concealment cost, the small country will experience a transition period of revenue loss. Its tax revenue will be higher than the initial level after countries adjust their tax rates. This implies that it has to take some time until the benefits of the GMT realize for all countries.

The rationale of part (i) is the following. Consider an increase in the GMT rate from t_2^N .

When $t_m \leq t_1^*$, country 1 chooses its tax rate along the initial best-response function. Thus, the revenue effect of the GMT rate for country 1 via the adjustment of its equilibrium tax rate cancels out (by the envelope theorem). On the other hand, a higher GMT rate creates a positive externality on country 1, since it raises the tax rate on affiliate 2' GloBE income and reduces the outward profit shifting in country 1. When $t_m > t_1^*$, country 1 chooses between undercutting the minimum and setting a higher rate $t_1(t_m)$, depending on which tax revenue is larger. In this case, the revenue effect of increasing the minimum rate is no less than that in the case $t_m \leq t_1^*$. So introducing the GMT always benefits country 1.

In contrast, the revenue effect for country 2 is more delicate. Part (ii) states that a marginal reform of the GMT increases country 2's long-run revenue either when the

market-size asymmetry is *large* or when the profit shifting cost is *low*. The rationale is the following. Consider a marginal increase in the GMT rate from country 2's initial tax rate. For $t_2^N < t_2^*$, the GMT binds country 2 and induces country 1 to set tax rate along the initial best response curve. In this case, our result is the same as the traditional minimum taxation literature which treats the minimum tax as a lower bound imposed on countries' tax rates (see Kanbur and Keen, 1993; Keen and Konrad, 2013; Hebous and Keen, 2023). Absent the GMT country 2's revenue is maximized at t_2^N , which means that the marginal reform has a zero first-order effect on its equilibrium revenue (i.e., $\left. \frac{\partial R_2(t_m, t_1(t_m))}{\partial t_2} \right|_{t_m=t_2^N} = 0$). But it induces country 1 to increase tax rate due to the strategic complementarity (i.e., $t_1'(t_m) > 0$ by (A1)), which in turn imposes a positive externality on country 2's revenue. For $t_2^* \leq t_2^N \leq t_1^*$, country 2 undercuts the minimum and promotes the real investment to gain a higher tax revenue, while country 1 still chooses its best response $t_1(t_m)$. Consequently, the revenue effect of the marginal reform for country 2 is no less than that in the case $t_2^N < t_2^*$. In summary, the marginal GMT reform benefits country 2 when $t_2^N \leq t_1^*$ (or equivalently, by Lemma 4, when $\alpha_2 \leq \alpha_2^*$ or when $\alpha_2 > \alpha_2^*$ and $\delta \leq \delta^{**}$).

However, the above argument does not necessarily hold for $t_2^N > t_1^*$ (or equivalently, by Lemma 4, for $\alpha_2 > \alpha_2^*$ and $\delta > \delta^{**}$). Indeed, when the asymmetry between two countries is *small* and the profit shifting cost is *intermediate*, even a marginal reform of the GMT under Pillar Two may harm the low-tax country. This is substantially different from the previous minimum taxation literature. To explain the underlying intuition, we rewrite country 2's equilibrium revenue as follows:

$$R_2^m(t_2^m, t_1^m) = \underbrace{t_m (f_2 - \mu r k_2)}_{\text{taxation on true profit}} - \underbrace{(t_m - t_2^m) \sigma k_2}_{\text{taxation on shifted profit}} + \underbrace{t_m g^m(t_m)}_{\text{taxation on shifted profit}}.$$

When $t_2^N > t_1^*$ and $R_1^m(\tilde{t}_1, \tilde{t}_2)|_{t_m=t_2^N} > R_1(t_1^N, t_2^N)$, a marginal increase in the GMT rate from t_2^N induces both countries to undercut the minimum. The marginal reform increases the taxation on the true profit generated by substantive activities. However, the true tax rate differential vanishes so that the profit shifted to country 2 jumps *discontinuously* from

$\frac{t_1^N - t_2^N}{\delta}$ to zero. This is distinct from the traditional minimum taxation literature in which the

minimum rate binds the small country so that the profit shifting level changes continuously with the minimum rate. When the concealment cost δ is not very large, country 2 attracts considerable paper profits absent the GMT. Then the revenue loss from eliminating profit shifting can dominate the increased taxation on true profits, so that country 2's revenue decreases after introducing the marginal tax reform.

Part (iii) provides (sufficient) conditions for the Pareto-improving non-marginal reform. The intuition is straightforward. For $\sigma \geq \sigma_2^m$, the equilibrium investment level is interior such that country 2's revenue from taxing the true profit is larger than that without the GMT. For $t_m \leq t_1^*$, the GloBE incomes of affiliate 1 and 2 are taxed at $t_1(t_m)$ and t_m , respectively. It rules out the case where both countries undercut the minimum so that profit shifting jumps discretely to zero. When the profit shifting is not sensitive to the minimum rate changes (i.e., $\varepsilon_g(t_m) < 1$), the increase in the minimum rate outweighs the decrease in the inward profit shifting so that country 2's revenue from taxing the shifted profit also increases.

3.5. Conclusion

This paper has studied how the global minimum tax under Pillar Two of the OECD/G20 Inclusive Framework affects MNEs' behavior and countries' corporate taxes. To the best of our knowledge, our paper is the first to consider partial tax deductibility, the SBIE and the QDMTT simultaneously in a formal model of tax competition between asymmetric countries. In response to the tax policies, the MNE chooses capital investments and profit shifting to maximize total after-tax profits. Under the QDMTT, the country with tax rate below the minimum will collect top-up taxes from the MNE's affiliate operating in its territory.

Introducing a GMT with minimum rate lying between two countries' initial tax rates can reduce the true tax differentials, curb profit shifting and always raise the large country's tax revenue. The revenue effect for the small country, however, is generally ambiguous. In the short run where countries' tax rates are fixed, due to tax deduction of the SBIE, a higher minimum rate has two opposite effects. On one hand, it means that more taxes can be saved by the low-tax affiliate, which incentivizes the capital investment. On the other hand, it raises the top-up tax rate and thus incurs a larger revenue loss for the small country. We show that the former (latter) effect dominates under high (low) profit shifting cost so that the small country's tax revenue increases (decreases). In the long run where countries can adjust their

tax rates, the GMT reshapes the underlying tax game and the competition pattern. At the equilibrium, the minimum rate binds the small country only if it is low. Otherwise, the small country will set its tax rate below the minimum, even at zero (if the carve-out rate is very small), in order to promote real investments in its territory. When the GMT rate is sufficiently high and the carve-out is not very small, both countries undercut the minimum and so profit shifting vanishes. In this case, countries aim at attracting capital investments instead of competing for paper profits. Moreover, we show that countries' equilibrium tax rates are decreasing in the minimum rate whenever they are below the minimum. While the equilibrium tax rate of the small country changes continuously with the minimum, gradual increases in the GMT rate may trigger the large country's tax rate to drop discontinuously below the minimum. As for the long-run revenue effect, a marginal GMT reform does not necessarily benefit the small country. It can raise the small country's tax revenue if the market-size asymmetry is large or if the profit shifting cost is low. For small market-size asymmetry and intermediate profit shifting cost, the revenue loss from the elimination of profit shifting may outweigh the revenue gain from taxing the true profits generated by genuine economic activities, so that even a marginal reform may harm the small country. Our results are substantially different from the traditional minimum taxation literature which specifies the minimum tax as a lower bound imposed on countries' tax rates. We also investigate the Pareto-improving non-marginal tax reform. When the carve-out rate is not too small, the GMT rate is not very high and profit shifting is not very sensitive to the minimum rate, the GMT can benefit the small country by increasing the taxation of both true profit and shifted profit.

Our findings suggest that the introduction of the GMT may not bring all countries' tax rates above the minimum rate of 15%. Small countries may still set their taxes below the minimum to attract inward investments and collect top-taxes. Indeed, Ireland has decided to keep the corporate tax rate of 12.5% and top up the rate to 15% for Irish affiliates of MNEs (see Department of Finance, 2023). Moreover, introducing the GMT in low concealment cost environments – which might be the current situation due to cross-border tax loopholes – will cause a short period of revenue losses for small countries. But they can benefit from a moderate GMT rate in the long run where countries' tax rates are reset. This implies that it has to take some time until the benefits of the GMT realize for all countries.

Appendix 3.A. Proofs and derivations

3.A.1. Proof of Lemma 1(i)

Recall from (3.1) that one country's capital investment is zero when it sets a very high tax rate. From (3.2), the high-tax affiliate will report zero profit when the tax rate differential is large or profit shifting cost is very small. The possible corner solutions undermine the smoothness of tax revenue function and complicate the analysis. Our proof will handle this issue and show that the Nash equilibrium must be interior. We proceed in five steps.

Step 1. We show that $\pi_i(t_i^N, t_j^N) > 0$, $i = 1, 2$.

Given country j 's tax rate, country i can set its tax rate marginally above zero to earn a positive tax revenue. This implies that at equilibrium each country's revenue is positive. So the GloBE income (i.e., taxable profit) of affiliate i is

$$\pi_i(t_i^N, t_j^N) = f_i(k_i^N) - \mu r k_i^N - \frac{t_i^N - t_j^N}{\delta} > 0.$$

Step 2. We claim that $t_1^N < \frac{\alpha_1 - r}{\alpha_1 - \mu r}$.

If $t_1^N \geq \frac{\alpha_1 - r}{\alpha_1 - \mu r}$, then the capital investment in country 1 is $k_1 = 0$. Besides, in this case no profit is shifted to country 1, irrespective of country 2's tax rate. So country 1's tax base is 0, which contradicts $\pi_1(t_1^N, t_2^N) > 0$. Hence, it must be that $t_1^N < \frac{\alpha_1 - r}{\alpha_1 - \mu r}$.

Step 3. We claim that $t_2^N \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$.

To prove this claim, we analyze the following two cases, respectively. Case (i):

$t_1^N \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. In this case, $\forall t_2 \geq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$, $\pi_2(t_2, t_1^N) = 0 \Rightarrow t_2^N < \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. Case (ii):

$t_1^N > \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. In this case, $\forall t_2 \geq t_1^N$, $\pi_2(t_2, t_1^N) = 0 \Rightarrow t_2^N < t_1^N$. On the other hand, if

$\frac{\alpha_2 - r}{\alpha_2 - \mu r} < t_2^N < t_1^N$, then country 2's equilibrium revenue is $R_2(t_2^N, t_1^N) = \frac{t_2^N (t_1^N - t_2^N)}{\delta}$. Using

$\frac{\alpha_1 - r}{\alpha_1 - \mu r} \leq 2 \frac{\alpha_2 - r}{\alpha_2 - \mu r}$ and $t_1^N < \frac{\alpha_1 - r}{\alpha_1 - \mu r}$, we can derive: $\frac{t_1^N}{2} < \frac{\alpha_1 - r}{2(\alpha_1 - \mu r)} \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r} \Rightarrow$

$\frac{t_2(t_1^N - t_2)}{\delta}$ strictly decreases with t_2 for all $t_2 \geq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. Consider a tax rate t_2' that is

marginally below t_2^N . By the continuity of π_1 in t_2 , we get: $\pi_1(t_1^N, t_2') > 0$ and $k_2(t_1') = 0$.

Then $R_2(t_2', t_1^N) = \frac{t_2'(t_1^N - t_2')}{\delta} > R_2(t_2^N, t_1^N)$, which contradicts the definition of Nash

equilibrium. So it must be that $t_2^N \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$.

Step 4. We derive country i 's best response to t_j^N .

Using (3.1) – (3.3), we obtain that:

$$\forall (t_i, t_j) \in A_{ij} := \left\{ (t_i, t_j) \in \left[0, \frac{\alpha_i - r}{\alpha_i - \mu r} \right] \times \left[0, \frac{\alpha_j - r}{\alpha_j - \mu r} \right] : \pi_i(t_i, t_j) > 0 \text{ and } \pi_j(t_j, t_i) > 0 \right\}, \text{ country}$$

i 's tax revenue $R_i(t_i, t_j) = \varphi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta}$, where

$$\varphi_i(t_i) := t_i (f_i - \mu r k_i) = t_i \left[\frac{\alpha_i^2}{2} - \alpha_i \mu r - \frac{r^2 (1 - \mu t_i)(1 - 2\mu + \mu t_i)}{2(1 - t_i)^2} \right] \text{ is the revenue from taxing}$$

affiliate i 's true profit and $\frac{t_i(t_j - t_i)}{\delta}$ is the revenue gain (loss) from the MNE's profit

shifting if $t_j > t_i$ ($t_j < t_i$). From the results shown in Steps 1-3, we have $(t_i^N, t_j^N) \in A_{ij}$.

Besides, it is straightforward to show that: for all $t_j \in \left[0, \frac{\alpha_1 - r}{\alpha_1 - \mu r} \right]$,

$$\frac{\partial}{\partial t_i} \left[\varphi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta} \right] \Big|_{t_i=0} = \frac{(\alpha_i - r)(\alpha_i + r - 2\mu r)}{2} + \frac{t_j}{\delta} > 0,$$

$$\frac{\partial}{\partial t_i} \left[\varphi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta} \right] \Big|_{t_i = \frac{\alpha_i - r}{\alpha_i - \mu r}} = -\frac{(\alpha_i - r)(\alpha_i - \mu r)^2}{(1 - \mu)r} - \frac{\frac{2(\alpha_i - r)}{\alpha_i - \mu r} - t_j}{\delta} < 0, \text{ and}$$

$$\frac{\partial^2}{\partial t_i^2} \left[\varphi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta} \right] = -\frac{r^2 (1 - \mu)^2 (2 + t_i)}{(1 - t_i)^4} - \frac{2}{\delta} < 0, \text{ which leads to the following:}$$

(i) $\forall t_j \in \left[0, \frac{\alpha_1 - r}{\alpha_1 - \mu r}\right]$, $\frac{\partial}{\partial t_i} \left[\varphi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta} \right] = 0$ has a unique solution denoted by $t_i(t_j)$ with $t_i(t_j) \in \left(0, \frac{\alpha_i - r}{\alpha_i - \mu r}\right)$; and (ii) $\varphi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta}$ strictly increases (decreases) with t_i if $t_i < t_i(t_j)$ ($t_i > t_i(t_j)$).

Now we argue by contradiction that $t_i^N = t_i(t_j^N)$. If $t_i^N < t_i(t_j^N)$ ($t_i^N > t_i(t_j^N)$), then consider a tax rate t_i' that is marginally above (below) t_i^N . By the continuity of π_i and π_j in t_i , we would get: $\pi_i(t_i', t_j^N) > 0$ and $\pi_j(t_j^N, t_i') > 0$. Together with $t_i' \in \left(0, \frac{\alpha_i - r}{\alpha_i - \mu r}\right)$, it would be that: $(t_i', t_j^N) \in A_{ij} \Rightarrow R_i(t_i', t_j^N) > R_i(t_i^N, t_j^N)$, which contradicts the definition of Nash equilibrium. Therefore, it must be that $t_i^N = t_i(t_j^N)$.

Step 5. We establish the existence and uniqueness of the Nash equilibrium.

Given $\frac{\partial}{\partial t_i} \left[\varphi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta} \right] = \varphi_i'(t_i) + \frac{t_j - 2t_i}{\delta} = 0$, using the implicit function theorem yields: $\frac{dt_i(t_j)}{dt_j} = \frac{1}{2 - \delta\varphi_i''(t_i)} \in \left(0, \frac{1}{2}\right)$, (A1)

where $\varphi_i''(t_i) = -\frac{r^2(1-\mu)^2(2+t_i)}{(1-t_i)^4} < 0$.

(A1) shows that for interior solutions two countries' tax rates are strategic complements. Moreover, it implies that function $T(t_1, t_2) := (t_1(t_2), t_2(t_1))$ is a contraction on the space

$\left[0, \frac{\alpha_1 - r}{\alpha_1 - \mu r}\right] \times \left[0, \frac{\alpha_2 - r}{\alpha_2 - \mu r}\right]$. Then by the contraction mapping theorem, there exists a unique

fixed point of $T(t_1, t_2)$, that is, a unique Nash equilibrium with $t_i^N = t_i(t_j^N) \in \left(0, \frac{\alpha_i - r}{\alpha_i - \mu r}\right)$,

$i = 1, 2$. **Q.E.D.**

3.A.2. The comparative static results

By the continuity of π_i and π_j , together with $t_i^N \in \left(0, \frac{\alpha_i - r}{\alpha_i - \mu r}\right)$, we can show that: in

a small neighborhood of (t_i^N, t_j^N) , $(t_i, t_j) \in A_{ij}$ and $R_i(t_i, t_j) = \phi_i(t_i) + \frac{t_i(t_j - t_i)}{\delta}$. Then

totally differentiating $\frac{\partial R_i}{\partial t_i} \Big|_{(t_i^N, t_j^N)} = 0$ yields the comparative statics without the GMT:

$$\frac{\partial t_i^N}{\partial \alpha_i} = \frac{\alpha_i - \mu r}{|J|} \cdot \left[\frac{r^2 (1 - \mu)^2 (2 + t_j^N)}{(1 - t_j^N)^4} + \frac{2}{\delta} \right] > 0, \quad (\text{A2})$$

$$\frac{\partial t_j^N}{\partial \alpha_i} = \frac{\alpha_i - \mu r}{\delta |J|} > 0, \quad (\text{A3})$$

$$\frac{\partial t_i^N}{\partial \delta} = \frac{1}{|J|} \cdot \left[\frac{(1 - \mu)^2 r^2 (2t_i^N - t_j^N)(2 + t_j^N)}{(1 - t_j^N)^4 \delta^2} + \frac{3t_i^N}{\delta^3} \right], \quad (\text{A4})$$

where $|J| := \frac{\partial^2 R_1}{\partial t_1^2} \cdot \frac{\partial^2 R_2}{\partial t_2^2} - \frac{\partial^2 R_1}{\partial t_1 \partial t_2} \cdot \frac{\partial^2 R_2}{\partial t_1 \partial t_2} = \prod_{i=1}^2 \left[\frac{r^2 (1 - \mu)^2 (2 + t_i^N)}{(1 - t_i^N)^4} + \frac{2}{\delta} \right] - \frac{1}{\delta^2} > 0$ is the

Jacobian determinant of the system $\frac{\partial R_i}{\partial t_i}$ ($i = 1, 2$) evaluated at (t_1^N, t_2^N) .

3.A.3. Proof of Lemma 1(ii)

For identical countries with $\alpha_1 = \alpha_2$, it must be that $t_1^N = t_2^N$. On the other hand, using

(A2) and (A3) we can get: $\frac{\partial(t_1^N - t_2^N)}{\partial \alpha_1} = \frac{\alpha_1 - \mu r}{|J|} \cdot \left[\frac{r^2 (1 - \mu)^2 (2 + t_2)}{(1 - t_2)^4} + \frac{1}{\delta} \right] > 0$, which means

that $t_1^N(\alpha_1, \alpha_2) - t_2^N(\alpha_1, \alpha_2)$ is strictly increasing in α_1 . Hence, we can derive that:

$$\forall \alpha_1 > \alpha_2, \quad t_1^N(\alpha_1, \alpha_2) - t_2^N(\alpha_1, \alpha_2) > t_1^N(\alpha_2, \alpha_2) - t_2^N(\alpha_2, \alpha_2) = 0 \quad \Rightarrow \quad \forall \alpha_1 > \alpha_2, \quad t_1^N > t_2^N.$$

Q.E.D.

3.A.4. Proof of Proposition 1

It follows from (3.8) that $g^m = \frac{t_1^N - t_m}{\delta} < \frac{t_1^N - t_2^N}{\delta} = g^N$. Together with (3.3), (3.6) and

(3.9), we have $R_1^m > R_1^N$.

Now we prove (ii). Differentiating (3.7) with respect to t_m and evaluating it at t_2^N gives:

$$\left. \frac{\partial k_2^m}{\partial t_m} \right|_{t_m=t_2^N} = \frac{\partial k_2^N}{\partial t_2^N} + \underbrace{\frac{\sigma}{-(1-t_2^N)f_2''}}_{\text{tax incentive effect}}, \quad (\text{A5})$$

where $\frac{\partial k_2^N}{\partial t_2^N} = \frac{f_2' - \mu r}{(1-t_2^N)f_2''}$ is the derivative of k_2 evaluated at t_2^N in the absence of the GMT.

Differentiating (3.10) with respect to t_m and using (A5) yields:

$$\begin{aligned} \left. \frac{\partial R_2^m}{\partial t_m} \right|_{t_m=t_2^N} &= \left[f_2(k_2^m) - \mu r k_2^m + g^m + t_m \left((f_2' - \mu r) \frac{\partial k_2^m}{\partial t_m} + \frac{\partial g^m}{\partial t_m} \right) - \sigma k_2^m - (t_m - t_2^N) \sigma \frac{\partial k_2^m}{\partial t_m} \right] \Big|_{t_m=t_2^N} \\ &= \underbrace{f_2(k_2^N) - \mu r k_2^N + \frac{t_1^N - t_2^N}{\delta} + t_2^N \left((f_2' - \mu r) \frac{\partial k_2^N}{\partial t_2^N} - \frac{1}{\delta} \right)}_{\frac{\partial R_2^N}{\partial t_2^N} = 0} + \underbrace{\frac{t_2^N (f_2' - \mu r) \sigma}{-(1-t_2^N)f_2''}}_{\text{gain from tax incentive}} - \underbrace{\sigma k_2^N}_{\text{loss from SBIE}}. \\ &= - \frac{\sigma \left[\alpha_2 (1-t_2^N)^2 - r (1-2\mu t_2^N + \mu (t_2^N)^2) \right]}{(1-t_2^N)^2} \end{aligned} \quad (\text{A6})$$

Using (A6), it is straightforward to show:

$$\left. \frac{\partial R_2^m}{\partial t_m} \right|_{t_m=t_2^N} \geq 0 \Leftrightarrow t_2^N \geq 1 - \sqrt{\frac{r(1-\mu)}{\alpha_2 - \mu r}} =: t_2^*. \quad (\text{A7})$$

It immediately follows from (A7) that: for $t_2^N > t_2^*$ ($t_2^N < t_2^*$), $R_2^m(t_m) > R_2^m(t_2^N) = R_2^N$ ($R_2^m(t_m) < R_2^m(t_2^N) = R_2^N$) when t_m is marginally higher than t_2^N .

Lastly, we prove part (iii). Notice that for the continuous function $R_2^m(t_m)$, it is quasiconcave in t_m if and only if one of the following conditions holds: (i) $R_2^m(t_m)$ is increasing; (ii) $R_2^m(t_m)$ is decreasing; or (iii) there is some $t_m' \in (t_2^N, t_1^N)$ such that $R_2^m(t_m)$ is increasing for all $t_m \in (t_2^N, t_m')$ and is decreasing for all $t_m \in (t_m', t_1^N)$. On the other hand, recall from (A7) that $\left. \frac{\partial R_2^m}{\partial t_m} \right|_{t_m=t_2^N} < 0$ if $t_2^N < t_2^*$. Assuming that $R_2^m(t_m)$ is quasiconcave in t_m for all $t_m \in (t_2^N, t_1^N)$, we can derive that: $R_2^m(t_m)$ must be decreasing in t_m if $t_2^N < t_2^*$ $\Rightarrow R_2^m(t_m) < R_2^m(t_2^N) = R_2^N$ holds for all $t_m \in (t_2^N, t_1^N)$, if $t_2^N < t_2^*$. **Q.E.D.**

3.A.5. Proof of Lemma 2

One might hope to prove the lemma by comparing $t_2^N(\delta)$ and t_2^* as $\delta \rightarrow 0$ and $\delta \rightarrow +\infty$, and then by resorting to the monotonicity of $t_2^N(\delta)$. However, unlike $t_1^N(\delta)$, $t_2^N(\delta)$ is not necessarily increasing in δ . To surmount this difficulty, we will invoke the Poincaré-Hopf index theorem (see Vives, 2000), which only requires information on $\frac{dt_2^N(\delta)}{d\delta}$ at δ that satisfies $t_2^N(\delta) = t_2^*$.

Define $\xi(\delta) := t_2^N(\delta) - t_2^*$. Recall from Appendix 3.A.1 that the equilibrium tax rates have to satisfy $\varphi_i'(t_i^N) - \frac{2t_i^N - t_j^N}{\delta} = 0$ ($i, j \in \{1, 2\}$ and $i \neq j$). Then it is straightforward to show the following:

$$\lim_{\delta \rightarrow 0} t_i^N(\delta) = 0, \quad i = 1, 2 \quad \Rightarrow \quad \lim_{\delta \rightarrow 0} \xi(\delta) < 0. \quad (\text{A8})$$

$$\phi'_2(t_2^*) = \frac{1}{2}(\alpha_2 - \mu r) \left[(\alpha_2 - \mu r) + (1 - \mu)r - 2\sqrt{(\alpha_2 - \mu r)(1 - \mu)r} \right] > 0 = \phi'_2(\lim_{\delta \rightarrow +\infty} t_2^N(\delta)) \Rightarrow$$

$$\lim_{\delta \rightarrow +\infty} t_2^N(\delta) > t_2^* \Rightarrow \lim_{\delta \rightarrow +\infty} \xi(\delta) > 0. \quad (\text{A9})$$

For all δ satisfying $t_2^N(\delta) = t_2^*$, using (A4) we can derive that:

$$\frac{2t_2^N(\delta) - t_1^N(\delta)}{\delta} = \phi'_2(t_2^N(\delta)) = \phi'_2(t_2^*) > 0 \Rightarrow \frac{dt_2^N(\delta)}{d\delta} > 0 \Rightarrow \xi'(\delta) > 0. \quad (\text{A10})$$

(A10) indicates that $\xi'(\delta)$ is positive whenever $\xi(\delta) = 0$. Together with (A8) and (A9) (the boundary conditions), it follows from the Poincaré-Hopf index theorem that there exists a unique δ^* such that $\xi(\delta^*) = 0$. Furthermore, given (A8), (A9) and the uniqueness of δ^* , it is straightforward to show $t_2^N \geq t_2^* \Leftrightarrow \delta \geq \delta^*$. **Q.E.D.**

3.A.6. Proof of Lemma 3

Denote $\hat{t}_i := \min \left\{ \frac{\alpha_i(1 - t_m) - (1 - \mu t_m)r + \sigma t_m}{\sigma}, t_m \right\}$. Given (3.12), (3.14) and (3.15), for

$t_i \in [\hat{t}_i, t_m)$, there is no investment in country i , and its revenue is $R_i^m = t_m(-1)^i g^m$, which is independent of t_i . For $t_i \in [0, \hat{t}_i)$, inserting (3.12) and (3.14) into (3.15) and differentiating the revenue function with respect to t_i yields:

$$\frac{\partial R_i^m}{\partial t_i} = t_m(f'_i - \mu r) \frac{\partial k_i}{\partial t_i} + \sigma k_i - (t_m - t_i)\sigma \frac{\partial k_i}{\partial t_i} = \frac{\sigma \left[\alpha_i(1 - t_m)^2 - r(1 - 2\mu t_m + \mu t_m^2) + (t_m - t_i)(2 - t_m)\sigma \right]}{(1 - t_m)^2}.$$

Moreover, it is readily verified that $\forall t_i \in [0, \hat{t}_i)$, $\frac{\partial^2 R_i^m}{\partial t_i^2} = -\frac{(2 - t_m)\sigma^2}{(1 - t_m)^2} < 0$. Then it is

straightforward to show the following:

(i) For $t_m \leq 1 - \sqrt{\frac{(1-\mu)r}{\alpha_i - \mu r}} =: t_i^*$, $\frac{\partial R_i^m}{\partial t_i} > 0$, $\forall t_i \in [0, t_m)$. That is, country i 's revenue

strictly increases with its tax rate. Besides, we have $R_i^m(t_m, t_j) = \varphi_i(t_m) + t_m(-1)^i g^m$.

(ii) For $t_m > t_i^*$ and $\underline{\sigma} < \sigma \leq \frac{r(1-2\mu t_m + \mu t_m^2) - \alpha_i(1-t_m)^2}{t_m(2-t_m)} =: \sigma_i^m$, $\forall t_i \in [0, \hat{t}_i)$,

$\frac{\partial R_i^m}{\partial t_i} < 0$. That is, country i 's revenue strictly decreases with t_i when $t_i \in [0, \hat{t}_i)$ and

remains unchanged when $t_i \in [\hat{t}_i, t_m)$. So the revenue-maximizing tax rate is zero, with

$$R_i^m(0, t_j) = \frac{(\alpha_i - r)^2}{2(2-t_m)} - \frac{(\sigma_i^m - \sigma)^2(2-t_m)t_m^2}{2(1-t_m)^2} + t_m(-1)^i g^m.$$

(iii) For $t_m > t_i^*$ and $\max\{\underline{\sigma}, \sigma_i^m\} < \sigma \leq \bar{\sigma}$, solving $\frac{\partial R_i^m}{\partial t_i} = 0$ yields

$t_i = \left(1 - \frac{\sigma_i^m}{\sigma}\right)t_m \in (0, \hat{t}_i)$. Country i 's revenue strictly increases (decreases) with t_i when

$t_i \in \left[0, \left(1 - \frac{\sigma_i^m}{\sigma}\right)t_m\right)$ ($t_i \in \left(\left(1 - \frac{\sigma_i^m}{\sigma}\right)t_m, \hat{t}_i\right)$) and remains unchanged when $t_i \in [\hat{t}_i, t_m)$. So the

revenue-maximizing tax rate is $\left(1 - \frac{\sigma_i^m}{\sigma}\right)t_m$, with $k_i^m|_{t_i=(1-\frac{\sigma_i^m}{\sigma})t_m} = \frac{\alpha_i - r}{2-t_m}$ and

$$R_i^m\left(\left(1 - \frac{\sigma_i^m}{\sigma}\right)t_m, t_j\right) = \frac{(\alpha_i - r)^2}{2(2-t_m)} + t_m(-1)^i g^m. \quad \mathbf{Q.E.D.}$$

3.A.7. Proof of Proposition 2

As shown in the proof of Lemma 3, given country j 's tax rate, country i can choose

$t_i = \min\{\tilde{t}_i, t_m\}$ to earn a positive tax revenue, where $\tilde{t}_i := \max\left\{0, \left(1 - \frac{\sigma_i^m}{\sigma}\right)t_m\right\}$. This implies

that each country's equilibrium revenue is positive $\Rightarrow$ the GloBE income $\pi_i^m(t_i^m, t_j^m) > 0$,

$i = 1, 2$. Notice that $t_m < t_1^N < \frac{\alpha_1 - r}{\alpha_1 - \mu r}$, which means that the GMT is inactive for affiliate 1

when country 1's tax rate $t_1 \geq \frac{\alpha_1 - r}{\alpha_1 - \mu r}$. Then repeating Step 2 in the proof of Lemma 1 with

superscript N replaced by m , we can show $t_1^m < \frac{\alpha_1 - r}{\alpha_1 - \mu r}$. In the following, we prove the

four parts of Proposition 2, separately.

3.A.7.1. Proof of Proposition 2(i)

Given $t_m \leq t_2^* < t_1^*$, it follows from Lemma 3(i) that: for each country, any tax choice below the GMT rate is strictly dominated by t_m such that its equilibrium tax must meet $t_i^m \geq t_m$. This means that at equilibrium the GMT is inactive for both affiliates. Each

country's equilibrium tax base (GloBE income) is: $\pi_i^m(t_i^m, t_j^m) = f_i(k_i^m) - \mu r k_i^m - \frac{t_i^m - t_j^m}{\delta} > 0$.

Step 1. We claim that $t_2^m \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$.

Notice that $t_m \leq t_2^* \Rightarrow t_m < \frac{\alpha_2 - r}{\alpha_2 - \mu r}$ and that $R_i^m(t_i, t_j) = R_i(t_i, t_j)$ when $t_i \geq t_m$ and

$t_j \geq t_m$. Then repeating Step 3 in the proof of Lemma 1 with superscript N replaced by m ,

we have: $t_2^m \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$.

Step 2. We claim that $t_i^m = \max\{t_m, t_i(t_j^m)\}$.

We analyze two cases, respectively. Case (i): $t_m \geq t_i(t_j^m)$. We argue by contradiction that

$t_i^m = t_m$. If $t_i^m > t_m \geq t_i(t_j^m)$, then consider a tax rate t_i' that is marginally below t_i^m . Notice

that $(t_i^m, t_j^m) \in A_{ij}$. Then by the continuity of π_i and π_j in t_i , we have: $\pi_i(t_i', t_j^m) > 0$ and

$\pi_j(t_j^m, t_i') > 0 \Rightarrow (t_i', t_j^m) \in A_{ij} \Rightarrow R_i^m(t_i', t_j^m) = R_i(t_i', t_j^m) > R_i(t_i^m, t_j^m) = R_i^m(t_i^m, t_j^m)$, which contradicts the definition of Nash equilibrium. So it must be that $t_i^m = t_m$. Case (ii): $t_m < t_i(t_j^m)$. By similar reasoning as in Case (i), we can show $t_i^m = t_i(t_j^m)$. Combining the two cases, we have proved the claim.

Step 3. The Nash equilibrium taxes are $t_1^m = t_1(t_m)$, $t_2^m = t_m$.

Using (A1) and $t_i^m = \max\{t_m, t_i(t_j^m)\}$, we can derive: $t_1(t_2^m) \geq t_1(t_m) > t_1(t_2^N) = t_1^N > t_m \Rightarrow t_1^m = t_1(t_2^m)$. Now we argue that $t_2^m \neq t_2(t_1^m)$. If $t_2^m = t_2(t_1^m)$, then together with $t_1^m = t_1(t_2^m)$ we would obtain $t_2^m = t_2^N$, which contradicts $t_2^m \geq t_m > t_2^N$. So given $t_2^m = \max\{t_m, t_2(t_1^m)\}$, it must be that $t_2^m = t_m$. So the equilibrium taxes have to satisfy $t_1^m = t_1(t_m)$, $t_2^m = t_m$. On the other hand, it follows from (A1) that the function $t_2(t_1(x)) - x$ strictly decreases with x for all $x \in \left[0, \frac{\alpha_1 - r}{\alpha_1 - \mu r}\right]$. Then we can derive:

$$t_2(t_1^m) - t_m = t_2(t_1(t_m)) - t_m < t_2(t_1(t_2^N)) - t_2^N = 0 \Rightarrow t_2(t_1^m) < t_m \Rightarrow \max\{t_m, t_2(t_1^m)\} = t_m. \quad (\text{A11})$$

So country 2's best response to country 1's tax choice $t_1(t_m)$ is choosing t_m . Hence, $(t_1(t_m), t_m)$ satisfies the definition of Nash equilibrium. **Q.E.D.**

3.A.7.2. Proof of Proposition 2(ii)

Firstly, it follows from Lemma 3 (i) that: for country 1, any tax choice below the GMT rate is strictly dominated by t_m such that its equilibrium tax must meet $t_1^m \geq t_m$.

Step 1. We claim that $t_2^m = \tilde{t}_2$.

We analyze two cases: (i) $t_m \geq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$, and (ii) $t_m < \frac{\alpha_2 - r}{\alpha_2 - \mu r}$, separately.

Case (i): $t_m \geq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. In this case, we prove by contradiction that $t_2^m \leq t_m$. If $t_2^m > t_m$

and $t_1^m \leq t_2^m$, then we would have $\pi_2(t_2^m, t_1^m) = 0$, which contradicts $\pi_2(t_2^m, t_1^m) > 0$. If

$t_2^m > t_m$ and $t_1^m > t_2^m$, then country 2's equilibrium revenue would be

$R_2^m(t_2^m, t_1^m) = \frac{t_2^m(t_1^m - t_2^m)}{\delta}$. Then by the same reasoning as in Step 3 in the proof of Lemma 1,

for a tax rate t_2' that is marginally below t_2^m , we would get: $R_2^m(t_2', t_1^m) > R_2^m(t_2^m, t_1^m)$,

which contradicts the definition of Nash equilibrium. So we have shown $t_2^m \leq t_m$. Then it

follows from Lemma 3 that $t_2^m = \tilde{t}_2$.

Case (ii): $t_m < \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. We argue by contradiction that $t_2^m < t_m$. If $t_2^m \geq t_m$, then

repeating Steps 1 – 3 in the proof of Proposition 2(i) would lead to $t_1^m = t_1(t_m)$ and $t_2^m = t_m$.

However, it follows from Lemma 3 that $R_2^m(\tilde{t}_2, t_1^m) > R_2^m(t_m, t_1^m)$, which contradicts the

definition of the Nash equilibrium. So country 2's equilibrium tax must satisfy $t_2^m < t_m$. Then

using Lemma 3 again, it must be that $t_2^m = \tilde{t}_2$.

Step 2. Given $t_2^m = \tilde{t}_2$, country 1's best response is choosing $t_1(t_m)$.

Given $t_2^m = \tilde{t}_2$, country 1's revenue function is: $R_1^m(t_1, t_2^m) = \begin{cases} R_1(t_1, t_m) & \text{if } t_m \leq t_1 < \hat{t}_1 \\ 0 & \text{if } t_1 \geq \hat{t}_1 \end{cases}$,

where $\hat{t}_1 \in \left(t_1(t_m), \frac{\alpha_1 - r}{\alpha_1 - \mu r}\right)$ is the unique value of t_1 such that

$\pi_1^m(\hat{t}_1, t_2^m) = \left[f_1(k_1) - \mu r k_1 - \frac{t_1 - t_m}{\delta} \right] \Big|_{t_1 = \hat{t}_1} = 0$ (noticing that the expression for $R_1^m(t_1, t_2^m)$ holds

for all $t_2^m \in [0, t_m]$). Besides, using (A1) and $t_2^N < t_m < t_1^N$ yields $t_1(t_m) > t_1(t_2^N) = t_1^N > t_m$.

Hence, $R_1^m(t_1, t_2^m)$ increases (decreases) with t_1 if $t_m \leq t_1 < t_1(t_m)$ ($t_1 > t_1(t_m)$), such that country 1's best response is setting a tax rate of $t_1(t_m)$.

Step 3. We show that $(t_1(t_m), \tilde{t}_2)$ are Nash equilibrium tax rates.

Given the result shown in Step 2, we only need to analyze country 2's best response to $t_1^m = t_1(t_m)$. Given country 1's tax choice $t_1^m = t_1(t_m)$, country 2's revenue function depends on the relationship between t_m , t_1^m and $\frac{\alpha_2 - r}{\alpha_2 - \mu r}$ in the following way:

$$\text{When } \frac{\alpha_2 - r}{\alpha_2 - \mu r} \leq t_m < t_1^m, \quad R_2^m(t_2, t_1^m) = \begin{cases} \frac{t_2(t_1^m - t_2)}{\delta} & \text{if } t_m \leq t_2 \leq t_1^m \\ 0 & \text{if } t_2 > t_1^m \end{cases}; \quad \text{When } t_m < \frac{\alpha_2 - r}{\alpha_2 - \mu r} < t_1^m, \quad R_2^m(t_2, t_1^m) = \begin{cases} R_2(t_2, t_1^m) & \text{if } t_m \leq t_2 < \frac{\alpha_2 - r}{\alpha_2 - \mu r} \\ \frac{t_2(t_1^m - t_2)}{\delta} & \text{if } \frac{\alpha_2 - r}{\alpha_2 - \mu r} \leq t_2 < t_1^m \\ 0 & \text{if } t_2 \geq t_1^m \end{cases}; \quad \text{When } t_m < t_1^m \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r},$$

$$R_2^m(t_2, t_1^m) = \begin{cases} R_2(t_2, t_1^m) & \text{if } t_m \leq t_2 < \hat{t}_2 \\ 0 & \text{if } t_2 \geq \hat{t}_2 \end{cases}, \quad \text{where } \hat{t}_2 \in \left[t_1^m, \frac{\alpha_2 - r}{\alpha_2 - \mu r} \right] \text{ is the unique value of } t_2 \text{ such that } \pi_2^m(\hat{t}_2, t_1^m) = \left[f_2(k_2) - \mu r k_2 - \frac{t_2 - t_1^m}{\delta} \right] \Big|_{t_2 = \hat{t}_2} = 0. \text{ Note that } \frac{t_1^m}{2} < \frac{\alpha_1 - r}{2(\alpha_1 - \mu r)} \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r},$$

and recall from (A11) that $t_2(t_1^m) < t_m$. It is straightforward to show that $R_2^m(t_2, t_1^m)$ always decreases with t_2 for all $t_2 \geq t_m$. Together with Lemma 3, country 2's best response is choosing $\tilde{t}_2$. **Q.E.D.**

3.A.7.3. Proof of Proposition 2(iii)

We proceed in two steps.

Step 1. We claim that $t_2^m = \tilde{t}_2$.

If country 1's equilibrium tax rate $t_1^m \geq t_m$, then by repeating Step 1 in the proof of Proposition 2(ii), it must be that $t_2^m = \tilde{t}_2$.

If country 1's equilibrium tax $t_1^m < t_m$, then it follows from Lemma 3 that $t_1^m = \tilde{t}_1$. In what follows, we analyze two cases. Case (i): $t_m \geq \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. It is obvious that $R_2^m(t_2, t_1^m) = 0$

for all $t_2 \geq t_m$. Then using Lemma 3, we have $t_2^m = \tilde{t}_2$. Case (ii): $t_m < \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. For $t_2 \geq t_m$,

country 2's tax revenue function is $R_2^m(t_2, t_1^m) = \begin{cases} R_2(t_2, t_m) & \text{if } t_m \leq t_2 < \hat{t}_2 \\ 0 & \text{if } t_2 \geq \hat{t}_2 \end{cases}$, where

$\hat{t}_2 \in \left(t_m, \frac{\alpha_2 - r}{\alpha_2 - \mu r} \right)$ is the unique value of t_2 such that

$$\pi_2^m(\hat{t}_2, t_1^m) = \left[f_2(k_2) - \mu r k_2 - \frac{t_2 - t_m}{\delta} \right]_{t_2 = \hat{t}_2} = 0. \text{ Besides, it follows from (A1) and } t_2^N < t_m < t_1^N$$

that: $t_2(t_m) < t_2(t_1^N) = t_2^N < t_m$, which implies that $R_2^m(t_2, t_1^m)$ decreases with t_2 for all $t_2 \geq t_m$. Together with Lemma 3, country 2's best response to $t_1^m = \tilde{t}_1$ is choosing $\tilde{t}_2$.

Step 2. We derive country 1's best response to $t_2^m = \tilde{t}_2$.

Using Lemma 3 and repeating Step 2 in the proof of Proposition 2(ii), we can obtain: $\arg \max_{t_1 \in [0, t_m]} R_1^m(t_1, t_2^m) = \tilde{t}_1$ and $\arg \max_{t_1 \in [t_m, 1]} R_1^m(t_1, t_2^m) = t_1(t_m)$. Therefore, country 1's best response

is choosing $t_1(t_m)$ when $R_1^m(t_1(t_m), t_2^m) = R_1(t_1(t_m), t_m) > R_1^m(\tilde{t}_1, t_2^m)$, and $\tilde{t}_1$ when

$R_1^m(t_1(t_m), t_2^m) = R_1(t_1(t_m), t_m) < R_1^m(\tilde{t}_1, t_2^m)$. It is indifferent between choosing $t_1(t_m)$ and $\tilde{t}_1$

when $R_1^m(t_1(t_m), t_2^m) = R_1(t_1(t_m), t_m) = R_1^m(\tilde{t}_1, t_2^m)$. **Q.E.D.**

3.A.7.4. Proof of Proposition 2(iv)

We first prove $t^{**} \in (t_1^*, \bar{t}_1)$. Using $\frac{(\alpha_1 - r)^2}{2(2 - t^{**})} = \bar{R}_1 = \varphi_1(\bar{t}_1)$, we can show:

$$\frac{(\alpha_1 - r)^2}{2(2 - \bar{t}_1)} - \frac{(\alpha_1 - r)^2}{2(2 - t^{**})} = \frac{(\alpha_1 - r)^2}{2(2 - \bar{t}_1)} - \varphi_1(\bar{t}_1) = \frac{\left[\alpha_1(1 - \bar{t}_1)^2 - r(1 - 2\mu\bar{t}_1 + \mu\bar{t}_1^2) \right]^2}{2(2 - \bar{t}_1)(1 - \bar{t}_1)^2} > 0 \Rightarrow t^{**} < \bar{t}_1. \quad (\text{A12})$$

On the other hand, using $\varphi_1'(\bar{t}_1) = 0$ we can obtain:

$$\frac{\alpha_1}{r} = \sqrt{\frac{1 + \bar{t}_1}{(1 - \bar{t}_1)^3}} (1 - \mu) + \mu. \quad (\text{A13})$$

Using (A13), we have:

$$\bar{R}_1 = \varphi_1(\bar{t}_1) = r^2 \bar{t}_1 \left[\frac{\alpha_1^2}{2r^2} - \mu \frac{\alpha_1}{r} - \frac{(1 - \mu\bar{t}_1)(1 - 2\mu + \mu\bar{t}_1)}{2(1 - \bar{t}_1)^2} \right] = \frac{r^2(1 - \mu)^2 \bar{t}_1^2}{(1 - \bar{t}_1)^3}.$$

Then solving $\frac{(\alpha_1 - r)^2}{2(2 - t^{**})} = \bar{R}_1 = \frac{r^2(1 - \mu)^2 \bar{t}_1^2}{(1 - \bar{t}_1)^3}$ for t^{**} and using (A13), we can get:

$$t^{**} = 2 - \frac{(1 - \bar{t}_1)^3 \left(\frac{\alpha_1}{r} - 1 \right)^2}{2(1 - \mu)^2 \bar{t}_1^2} = 2 - \frac{(1 - \bar{t}_1)^3}{2\bar{t}_1^2} \cdot \left(\sqrt{\frac{1 + \bar{t}_1}{(1 - \bar{t}_1)^3}} - 1 \right)^2. \quad (\text{A14})$$

Define $H(t) := \frac{\alpha_1}{r}(1 - t)^2 - (1 - 2\mu t + \mu t^2)$. It is readily verified that $H(t) < 0$ if and only if $t \in \left(t_1^*, 1 + \sqrt{\frac{r(1 - \mu)}{\alpha_1 - \mu r}} \right)$. Plugging (A13) and (A14) into $H(t)$ yields:

$$H(t^{**}) = - \frac{(1 - \mu) \left(1 - \sqrt{1 - \bar{t}_1^2} \right)^3 \left(5 + 3\sqrt{1 - \bar{t}_1^2} - \bar{t}_1^2 \right)}{4\bar{t}_1^4} < 0 \Rightarrow t^{**} > t_1^*. \quad (\text{A15})$$

Combining (A12) and (A15) leads to $t^{**} \in (t_1^*, \bar{t}_1)$.

Notice that $\forall t_m \in (t_2^N, t_1^N)$, $\bar{t}_1 > t_1(t_m) > t_m$. Then replacing t_1^N with $t_1(t_m)$ and t_2^N with t_m in (3.4), we have:

$$R_1(t_1(t_m), t_m) < \bar{R}_1. \quad (\text{A16})$$

Let δ be an arbitrary element of set $\{\delta > 0 : t_1^N(\delta) > t^{**}\}$, which is nonempty according to (A12). When $\max\{t_2^N(\delta), t^{**}\} < t_m < t_1^N(\delta)$ and $\max\{\underline{\sigma}, \sigma_1^m\} < \sigma \leq \bar{\sigma}$, by (A15) and (A16) we can show that:

$$t_m > t_1^* \text{ and } R_1^m(\tilde{t}_1, \tilde{t}_2) = \frac{(\alpha_1 - r)^2}{2(2 - t_m)} > \frac{(\alpha_1 - r)^2}{2(2 - t^{**})} = \bar{R}_1 > R_1(t_1(t_m), t_m).$$

Then by Proposition 2(iii), the Nash equilibrium rates are $t_1^m = \tilde{t}_1$, $t_2^m = \tilde{t}_2$. **Q.E.D.**

3.A.8. Proof of Lemma 4

As in the proof of Lemma 2, we will again appeal to the Poincaré-Hopf index theorem, which only requires information on $\frac{dt_2^N(\delta)}{d\delta}$ at δ that satisfies $t_2^N(\delta) = t_1^*$. We proceed in two steps.

Step 1. We claim that: for all δ satisfying $t_2^N(\delta) = t_1^*$, $\frac{dt_2^N(\delta)}{d\delta} > 0$.

First, we prove by contradiction that $2t_2^N > t_1^N$ when $t_2^N = t_1^*$. Recall that equilibrium taxes without the GMT are determined by $\varphi'_i(t_i^N) - \frac{2t_i^N - t_j^N}{\delta} = 0$. If $2t_2^N \leq t_1^N$, then we would get:

$$\varphi'_1(2t_2^N) \geq \varphi'_1(t_1^N) = \frac{2t_1^N - t_2^N}{\delta} > 0 \geq \frac{2t_2^N - t_1^N}{\delta} = \varphi'_2(t_2^N), \quad (\text{A17})$$

where the first inequality holds since $\varphi'_1(t_1)$ strictly decreases with t_1 for all $t_1 \in [0, 1]$.

On the other hand, we could derive the following:

$$\begin{aligned}
\phi'_1(2t_2^N) - \phi'_2(t_2^N) &= \frac{1}{2} \left\{ (\alpha_1^2 - 2\mu r \alpha_1) - (\alpha_2^2 - 2\mu r \alpha_2) - \frac{r^2 t_2^N (1-\mu)^2 \left[4 - 9t_2^N + (t_2^N)^2 + (6t_2^N)^3 \right]}{(1-t_2^N)^3 (1-2t_2^N)^3} \right\} \\
&\leq \frac{1}{2} \left\{ (\alpha_1^2 - 2\mu r \alpha_1) - (\alpha_2^2 - 2\mu r \alpha_2) - \frac{r^2 t_2^N (1-\mu)^2 \left[4 - 9t_2^N + (t_2^N)^2 + (6t_2^N)^3 \right]}{(1-t_2^N)^3 (1-2t_2^N)^3} \right\} \\
&= \frac{r^2 (1-\mu)^2 t_2^N \Omega}{2(1-t_2^N)^4 (1-2t_2^N)^3 \left[2 - 2t_2^N + (t_2^N)^2 \right]^2} < 0,
\end{aligned}$$

where the inequality in line 2 holds since $\alpha_2^2 - 2\mu r \alpha_2$ strictly increases with α_2 for all

$\alpha_2 \geq \mu r$, the equality in line 3 is obtained by substitution $\alpha_1 = \frac{r \left[1 - 2\mu t_2^N + \mu (t_2^N)^2 \right]}{(1-t_2^N)^2}$ into

line 2, and

$$\Omega := 20t_2^N + 28(t_2^N)^2 - 151(t_2^N)^3 + 218(t_2^N)^4 - 167(t_2^N)^5 + 82(t_2^N)^6 - 29(t_2^N)^7 + 6(t_2^N)^8 - 8 < 0, \quad ,$$

$$\forall t_2^N \in (0, 1).$$

This contradicts (A17). So we have proved that $2t_2^N > t_1^N$ when $t_2^N = t_1^*$. Then using (A4), the claim directly follows.

Step 2. We prove Lemma 4.

Recall that $\phi'_2(t_2)$ strictly decreases with t_2 for all $t_2 \in [0, 1)$. Then we can derive the following:

$$t_1^* \geq \lim_{\delta \rightarrow +\infty} t_2^N(\delta) \Leftrightarrow \phi'_2(t_1^*) \leq \phi'_2(\lim_{\delta \rightarrow +\infty} t_2^N(\delta)) = 0 \Leftrightarrow \alpha_2 \leq \alpha_2^*,$$

$$\text{where } \alpha_2^* := \sqrt{(\alpha_1 - \mu r) \left[2\sqrt{r(1-\mu)(\alpha_1 - \mu r)} - r(1-\mu) \right]} + \mu r \in (\alpha_2, \alpha_1).$$

Define $\eta(\delta) := t_2^N(\delta) - t_1^*$. Then it follows from the claim shown in Step 1 that:

$$\eta'(\delta) > 0, \text{ whenever } \eta(\delta) = 0. \quad (\text{A18})$$

In what follows, we analyze three cases, respectively.

Case (i): $\alpha_2 > \alpha_2^*$. In this case, $\lim_{\delta \rightarrow +\infty} t_2^N(\delta) > t_1^*$. Recall that $\lim_{\delta \rightarrow 0} t_2^N(\delta) = 0 < t_1^*$. Together with (A18), by the same reasoning as in the proof of Lemma 2, we can show there exists a threshold δ^{**} such that $t_2^N \geq t_1^* \Leftrightarrow \delta \geq \delta^{**}$. In addition, by Lemma 2, we can derive that: $t_2^N(\delta^{**}) = t_1^* > t_2^* \Rightarrow \delta^{**} > \delta^*$.

Case (ii): $\alpha_2 < \alpha_2^*$. We prove by contradiction that $\forall \delta, t_2^N < t_1^*$. Assume that there were some $\delta_0 > 0$ such that $t_2^N(\delta_0) \geq t_1^*$. Notice that $\lim_{\delta \rightarrow +\infty} t_2^N(\delta) < t_1^*$. Then together with (A18), there would exist some δ' and δ'' with $\delta_0 \leq \delta' < \delta''$ such that $t_2^N(\delta') > t_1^*$ and $t_2^N(\delta'') < t_1^*$. So we would get: $-\eta(\delta') < 0$ and $-\eta(\delta'') > 0$ (the boundary conditions). By the Poincaré-Hopf index theorem, we would have:
$$\sum_{\delta \in \{\delta: \eta(\delta)=0 \text{ and } \delta' \leq \delta \leq \delta''\}} \text{sign}\{-\eta'(\delta)\} = +1,$$
 which contradicts (A18).

Case (iii): $\alpha_2 = \alpha_2^*$. Using the result shown in Case (ii), we get:

$$\forall \delta, t_2^N(\delta, \alpha_2^*) = \lim_{\alpha_2 \rightarrow (\alpha_2^*)^-} t_2^N(\delta, \alpha_2) \leq \lim_{\alpha_2 \rightarrow (\alpha_2^*)^-} t_1^* = t_1^*. \quad (\text{A19})$$

Moreover, note that $t_2^N(\delta, \alpha_2^*) \neq t_1^*$. Otherwise, using the claim shown in Step 1 we would get: $t_2^N(\delta', \alpha_2^*) > t_1^*$ when δ' is marginally larger than δ , which contradicts (A19). So it must be that $\forall \delta, t_2^N(\delta, \alpha_2^*) < t_1^*$. **Q.E.D.**

3.A.9. Proof of Proposition 3

By Proposition 2, we have: $R_1^m(t_1^m, t_2^m) = R_1(t_1(t_m), t_m)$ when $t_m \leq t_1^*$; and $R_1^m(t_1^m, t_2^m) = \max\{R_1(t_1(t_m), t_m), R_1^m(\tilde{t}_1, \tilde{t}_2)\} \geq R_1(t_1(t_m), t_m)$ when $t_m > t_1^*$. Besides, it is straightforward to show that: $\forall t_m \in (t_2^N, t_1^N)$, $\frac{dR_1(t_1(t_m), t_m)}{dt_m} = \frac{\partial R_1}{\partial t_2} = \frac{t_1(t_m)}{\delta} > 0 \Rightarrow R_1(t_1(t_m), t_m) > R_1(t_1(t_2^N), t_2^N) = R_1(t_1^N, t_2^N)$. Hence, the GMT always increases the large country's tax revenue.

Now we prove (ii). Firstly, we analyze the case of $t_2^N = t_1^*$. It is straightforward to show:

$$R_1(t_1(t_2^N), t_2^N) = R_1(t_1^N, t_2^N) > R_1(t_2^N, t_2^N) = \varphi_1(t_2^N) = \frac{(\alpha_1 - r)^2}{2(2 - t_2^N)}, \quad (\text{A20})$$

where the inequality holds by the definition of the Nash equilibrium and the last equality is obtained by using $t_2^N = t_1^*$.

Using (A20) and the continuity of $R_1(t_1(t_m), t_m)$ and $\frac{(\alpha_1 - r)^2}{2(2 - t_m)}$, together with

Proposition 2(iii), we can derive:

$$\left[R_1(t_1(t_m), t_m) - \frac{(\alpha_1 - r)^2}{2(2 - t_m)} \right]_{t_m = t_2^N} > 0 \quad \Rightarrow \quad \text{when } t_m \text{ is marginally above } t_2^N, \\ R_1(t_1(t_m), t_m) > \frac{(\alpha_1 - r)^2}{2(2 - t_m)} \geq R_1(\tilde{t}_1, \tilde{t}_2) \text{ and thus } (t_1^m, t_2^m) = (t_1(t_m), \tilde{t}_2). \quad (\text{A21})$$

Now consider introducing a GMT with minimum rate t_m marginally above t_2^N . It immediately follows from Proposition 2(i), (ii) and (A21) that: $R_2^m(t_2^m, t_1^m) = R_2(t_m, t_1(t_m))$ if $t_2^N < t_2^*$; and $R_2^m(t_2^m, t_1^m) = R_2^m(\tilde{t}_2, t_1(t_m)) > R_2(t_m, t_1(t_m))$ if $t_2^* \leq t_2^N \leq t_1^*$. Besides, using (A1) we can derive: $\frac{dR_2(t_m, t_1(t_m))}{dt_m} \Big|_{t_m = t_2^N} = \left[\frac{\partial R_2}{\partial t_1} \cdot \frac{dt_1(t_m)}{dt_m} \right]_{t_m = t_2^N} = \frac{t_2^N}{\delta} \cdot \frac{1}{2 - \delta \varphi_1'(t_1^N)} > 0 \Rightarrow$ when t_m is marginally above t_2^N , $R_2(t_m, t_1(t_m)) > R_2(t_2^N, t_1(t_2^N)) = R_2(t_2^N, t_1^N)$. So the marginal reform increases country 2's long-run revenue if $t_2^N \leq t_1^*$ (or equivalently, by Lemma 4, if $\alpha_2 \leq \alpha_2^*$, or if $\alpha_2 > \alpha_2^*$ and $\delta \leq \delta^{**}$).

On the other hand, numerical simulations indicate that: there exist parameters $(\alpha_1, \alpha_2, \mu, r, \delta, \sigma)$ satisfying (i) $t_2^N > t_1^*$, (ii) $\sigma_2^m(t_2^N) < \sigma < \bar{\sigma}(t_2^N)$, (iii)

$$\left[\frac{(\alpha_1 - r)^2}{2(2 - t_m)} - R_1(t_1(t_m), t_m) \right]_{t_m = t_2^N} = \frac{(\alpha_1 - r)^2}{2(2 - t_2^N)} - R_1(t_1^N, t_2^N) > 0, \quad \text{and} \quad (\text{iv})$$

$\left[\frac{(\alpha_2 - r)^2}{2(2 - t_m)} - R_2(t_1^N, t_2^N) \right]_{t_m = t_2^N} = \frac{(\alpha_2 - r)^2}{2(2 - t_2^N)} - R_2(t_1^N, t_2^N) < 0$. Then by the continuity of $\sigma_2^m(t_m)$,

$\bar{\sigma}(t_m)$, $R_1(t_1(t_m), t_m)$ and $\frac{(\alpha_i - r)^2}{2(2 - t_m)}$ ($i = 1, 2$) and using Proposition 2(iii), we can show the following:

When t_m is marginally above t_2^N , (i) $\sigma_2^m(t_m) < \sigma < \bar{\sigma}(t_m) \Rightarrow \tilde{t}_i = \left(1 - \frac{\sigma_i^m(t_m)}{\sigma}\right)t_m$,

$i = 1, 2$; (ii) $R_1^m(\tilde{t}_1, \tilde{t}_2) = \frac{(\alpha_1 - r)^2}{2(2 - t_m)} > R_1(t_1(t_m), t_m) \Rightarrow t_1^m = \tilde{t}_1$, $t_2^m = \tilde{t}_2$; and (iii)

$$R_2^m(\tilde{t}_2, \tilde{t}_1) = \frac{(\alpha_2 - r)^2}{2(2 - t_m)} < R_2(t_1^N, t_2^N).$$

Thus, for such parameters $(\alpha_1, \alpha_2, \mu, r, \delta, \sigma)$, introducing a marginal GMT reform induces both countries to undercut the minimum and reduces country 2's long-run revenue.

Lastly, we prove (iii). By Proposition 2, country 2's equilibrium tax revenue can be decomposed as follows:

$$R_2^m(t_2^m, t_1^m) = \underbrace{t_m(f_2 - \mu r k_2) - (t_m - t_2^m)\sigma k_2}_{\text{taxation on true profit}} + \underbrace{t_m g^m(t_m)}_{\text{taxation on shifted profit}},$$

$$\text{where } t_2^m = \begin{cases} t_m & \text{if } t_m \leq t_2^* \\ \tilde{t}_2 & \text{if } t_m > t_2^* \end{cases} \text{ and } g^m(t_m) = \begin{cases} 0 & \text{if } t_m > t_1^* \text{ and } R_1^m(\tilde{t}_1, \tilde{t}_2) > R_1(t_1(t_m), t_m) \\ \frac{t_1(t_m) - t_m}{\delta} & \text{otherwise} \end{cases}.$$

For $t_2^N < t_m \leq t_2^*$, recall that $\varphi_2'(t_2^*) > 0$ and $\varphi_2''(t_2) < 0$, which leads to: $t_m(f_2 - \mu r k_2) - (t_m - t_2^m)\sigma k_2 = \varphi_2(t_m) > \varphi_2(t_2^N)$. For $\sigma \in [\sigma_2^m, \bar{\sigma}]$ and $t_m > \max\{t_2^N, t_2^*\}$, it is straightforward to show: $t_m(f_2 - \mu r k_2) - (t_m - t_2^m)\sigma k_2 = \frac{(\alpha_2 - r)^2}{2(2 - t_m)} > \frac{(\alpha_2 - r)^2}{2(2 - t_2^N)} > \varphi_2(t_2^N)$.

Therefore, given that $\sigma \in [\sigma_2^m, \bar{\sigma}]$, the revenue from taxing the true profit is higher than that without the GMT.

On the other hand, when $t_m \leq t_1^*$, $\frac{\partial(t_m g^m)}{\partial t_m} = g^m(t_m)(1 - \varepsilon_g(t_m))$, where

$$\varepsilon_g(t_m) := -\frac{\partial g^m / g^m}{\partial t_m / t_m} = \frac{t_m(1 - t_1'(t_m))}{t_1(t_m) - t_m} > 0$$

is the elasticity of profit shifting with respect to the minimum rate (measured positively). Besides, using (A1) we can show that:

$$\frac{d\varepsilon_g(t_m)}{dt_m} = \frac{(1 - t_1'(t_m) - t_m t_1''(t_m))(t_1(t_m) - t_m) + t_m(1 - t_1'(t_m))^2}{(t_1(t_m) - t_m)^2} > 0, \quad \text{where}$$

$$t_1''(t_m) = \frac{\delta \varphi_1'''(t_1)}{[2 - \delta \varphi_1''(t_1)]^3} < 0 \quad \text{since} \quad \varphi_1'''(t_1) = -\frac{3r^2(3 + t_1)(1 - \mu)^2}{(1 - t_1)^5} < 0.$$

That is, $\varepsilon_g(t_m)$ strictly

increases with t_m . Then for any minimum t_m satisfying $t_m \leq t_1^*$ and $\varepsilon_g(t_m) \leq 1$, by the mean value theorem we have: $t_m g^m - t_2^N g_2^N = (t_m - t_2^N) \cdot g^m(\omega)[1 - \varepsilon_g(\omega)] > 0$, where $\omega \in (t_2^N, t_m)$. So the revenue from taxing the shifted profit is higher than that without the GMT.

In summary, the conditions for the GMT to increase the small country's revenue are: $\sigma \in [\sigma_2^m, \bar{\sigma}]$, $t_m \leq t_1^*$ and $\varepsilon_g(t_m) \in (0, 1]$. **Q.E.D.**

Appendix 3.B. Conditions for excess profits to be positive

3.B.1. The condition for short-run excess profit $E_2 > 0$

In the short run, the excess profit of affiliate 2 is: $E_2 = f_2(k_2^m) - \mu r k_2^m + g^m - \sigma k_2^m$, where k_2^m and g^m are given by (3.7) and (3.8), respectively. In what follows, we analyze two cases:

Case (i): $\sigma \leq \frac{(1 - \mu t_m)r - \alpha_2(1 - t_m)}{t_m - t_2^N}$. In this case, there is no investment in country 2, i.e.,

$$k_2^m = 0 \Rightarrow E_2 = g^m = \frac{t_1^N - t_m}{\delta} > 0.$$

Case (ii): $\sigma > \frac{(1-\mu t_m)r - \alpha_2(1-t_m)}{t_m - t_2^N}$. In this case,

$$k_2^m = \frac{\alpha_2(1-t_m) - (1-\mu t_m)r + (t_m - t_2^N)\sigma}{1-t_m} > 0 \quad \Rightarrow$$

$$E_2 = f_2(k_2^m) - \mu r k_2^m + g^m - \sigma k_2^m > f_2(k_2^m) - \mu r k_2^m - \sigma k_2^m = k_2^m \cdot \frac{\alpha_2(1-t_m) + r(1-(2-t_m)\mu) - (2-t_2^N - t_m)\sigma}{2(1-t_m)}$$

$$\Rightarrow E_2 > 0 \text{ if } \sigma \leq \frac{\alpha_2(1-t_m) + r(1-(2-t_m)\mu)}{2-t_2^N - t_m} =: \bar{\sigma}^S(t_m).$$

Combining Cases (i) and (ii), a sufficient condition for $E_2 > 0$ is: $\sigma \leq \bar{\sigma}^S(t_m)$.

3.B.2. The condition for long-run excess profit $E_i \geq 0$, $\forall t_i \in [0, t_m)$, $\forall i \in \{1, 2\}$.

Suppose that country i sets its tax rate below the minimum, i.e., $t_i < t_m$. Then we have:

$$E_i = f_i(k_i^m) - \mu r k_i^m + (-1)^i g^m - \sigma k_i^m \geq f_i(k_i^m) - \mu r k_i^m - \sigma k_i^m = \left(\alpha_i - \frac{k_i^m}{2} - \mu r - \sigma \right) k_i^m, \text{ where}$$

k_i^m and g^m are given by (3.12) and (3.14), respectively.

Given (3.12), k_i^m is decreasing in t_i with $k_i^m \geq 0$. So for all $t_i \in [0, t_m)$, $E_i \geq 0$ if

$$\alpha_i - \frac{k_i^m}{2} \Big|_{t_i=0} - \mu r - \sigma \geq 0 \quad (\text{or equivalently, if } \sigma \leq \frac{\alpha_i(1-t_m) + (1+\mu t_m - 2\mu)r}{2-t_m} =: \bar{\sigma}_i).$$

Therefore, $\forall t_i \in [0, t_m)$, $\forall i \in \{1, 2\}$, $E_i \geq 0$ if and only if $\sigma \leq \min\{\bar{\sigma}_1, \bar{\sigma}_2\} = \bar{\sigma}_2 =: \bar{\sigma}$.

Appendix 3.C. Sufficient conditions for the quasiconcavity of $R_2^m(t_m)$, $\forall t_m \in (t_2^N, t_1^N)$

Firstly, to ensure that the short-run excess profit $E_2 > 0$ for all $t_m \in (t_2^N, t_1^N)$, we impose an upper bound on the carve-out, i.e., $\sigma \leq \inf_{t_m \in (t_2^N, t_1^N)} \bar{\sigma}^S(t_m) = \bar{\sigma}^S(t_1^N)$. In what follows, we analyze two cases, separately.

Case (a): $\alpha_2(1-t_1^N) - (1-\mu t_1^N)r + (t_1^N - t_2^N)\sigma < 0$ (i.e., $\sigma < \frac{(\alpha_2 - \mu r)t_1^N - (\alpha_2 - r)}{t_1^N - t_2^N}$).

Given (3.7), we distinguish the following two subcases.

Subcase (a1): $t_m \in \left(t_2^N, \frac{\alpha_2 - r - \sigma t_2^N}{\alpha_2 - \mu r - \sigma}\right)$. We have $k_2^m = \frac{\alpha_2(1-t_m) - (1-\mu t_m)r + (t_m - t_2^N)\sigma}{1-t_m}$.

Together with (3.8) and (3.10), we can obtain:

$$\frac{d^2 R_2^m}{dt_m^2} = - \frac{\left[(1-t_2^N)\sigma - r(1-\mu) \right] \left[(2-t_2^N(4-t_m) + t_m)\sigma - r(2+t_m)(1-\mu) \right]}{(1-t_m)^4} - \frac{2}{\delta}. \quad (C1)$$

Besides, it is readily verified that $\frac{(\alpha_2 - \mu r)t_1^N - (\alpha_2 - r)}{t_1^N - t_2^N} < \frac{r(1-\mu)}{1-t_2^N} < \min \left\{ \frac{r(2+t_m)(1-\mu)}{2-t_2^N(4-t_m) + t_m}, \bar{\sigma}^s(t_1^N) \right\}$. Then it immediately

follows from (C1) that: $\forall t_m \in \left(t_2^N, \frac{\alpha_2 - r - \sigma t_2^N}{\alpha_2 - \mu r - \sigma}\right)$, $\frac{d^2 R_2^m}{dt_m^2} < 0$.

Subcase (a2): $t_m \in \left[\frac{\alpha_2 - r - \sigma t_2^N}{\alpha_2 - \mu r - \sigma}, t_1^N\right)$. We have $k_2^m = 0$ and $R_2^m = \frac{t_m(t_1^N - t_m)}{\delta}$.

Besides, we can derive: $\frac{t_1^N}{2} < \frac{\alpha_1 - r}{2(\alpha_1 - \mu r)} \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r} < \frac{\alpha_2 - r - \sigma t_2^N}{\alpha_2 - \mu r - \sigma} \Rightarrow R_2^m$ strictly

decreases with t_m for all $t_m \in \left[\frac{\alpha_2 - r - \sigma t_2^N}{\alpha_2 - \mu r - \sigma}, t_1^N\right)$.

Combining Subcases (a1) and (a2) indicates that: $R_2^m(t_m)$ is quasiconcave for

$t_m \in (t_2^N, t_1^N)$ when $\sigma < \frac{(\alpha_2 - \mu r)t_1^N - (\alpha_2 - r)}{t_1^N - t_2^N}$.

Case (b): $\alpha_2(1-t_1^N) - (1-\mu t_1^N)r + (t_1^N - t_2^N)\sigma \geq 0$ (i.e., $\sigma \geq \frac{(\alpha_2 - \mu r)t_1^N - (\alpha_2 - r)}{t_1^N - t_2^N}$).

From (3.6), for all $t_m \in (t_2^N, t_1^N)$, $k_2^m = \frac{\alpha_2(1-t_m) - (1-\mu t_m)r + (t_m - t_2^N)\sigma}{1-t_m}$ and $\frac{d^2 R_2^m}{dt_m^2}$ is

given by (C1). We distinguish two subcases.

Subcase (b1): $\sigma \in \left[\frac{(\alpha_2 - \mu r)t_1^N - (\alpha_2 - r)}{t_1^N - t_2^N}, \frac{r(1-\mu)}{1-t_2^N} \right]$. Recall that

$$\frac{r(1-\mu)}{1-t_2^N} < \min \left\{ \frac{r(2+t_m)(1-\mu)}{2-t_2^N(4-t_m)+t_m}, \bar{\sigma}^S(t_1^N) \right\}, \text{ from which we can show: } \forall t_m \in (t_2^N, t_1^N),$$

$$\frac{d^2 R_2^m}{dt_m^2} < 0.$$

Subcase (b2): $\sigma > \frac{r(1-\mu)}{1-t_2^N}$. It is straightforward to show: when

$$\sigma \geq \sup_{t_m \in (t_2^N, t_1^N)} \frac{r(2+t_m)(1-\mu)}{2-t_2^N(4-t_m)+t_m} = \frac{r(1-\mu)(2+t_2^N)}{(1-t_2^N)(2-t_2^N)}, \quad \frac{d^2 R_2^m}{dt_m^2} < 0 \text{ holds for all } t_m \in (t_2^N, t_1^N).$$

Combining Subcases (b1) and (b2) indicates that: $R_2^m(t_m)$ is concave (and so quasiconcave) for $t_m \in (t_2^N, t_1^N)$ when $\sigma \in \left[\frac{(\alpha_2 - \mu r)t_1^N - (\alpha_2 - r)}{t_1^N - t_2^N}, \frac{r(1-\mu)}{1-t_2^N} \right]$, or when

$$\sigma \in \left[\frac{r(1-\mu)(2+t_2^N)}{(1-t_2^N)(2-t_2^N)}, \bar{\sigma}^S(t_1^N) \right].$$

In summary, $R_2^m(t_m)$ is quasiconcave for $t_m \in (t_2^N, t_1^N)$, if $\sigma \leq \frac{r(1-\mu)}{1-t_2^N}$ or if

$$\frac{r(1-\mu)(2+t_2^N)}{(1-t_2^N)(2-t_2^N)} \leq \sigma \leq \bar{\sigma}^S(t_1^N).$$

Appendix 3.D. The analysis of the case of $\sigma \in (0, \underline{\sigma}]$

Notice that $0 < \sigma \leq \underline{\sigma} \Rightarrow t_m > \frac{\alpha_2 - r}{\alpha_2 - \mu r}$. Then given (3.1) and (3.12), it is

straightforward to show that: $\forall t_2 \in [0,1], k_2 = 0$. So when $\sigma \leq \underline{\sigma}$, country 2 becomes a tax haven with no capital investment such that the shifted profits (if any) constitute its only tax base. The following proposition states that there exists a continuum of Nash equilibria in this case.

Proposition D.1. *In the case of $\sigma \in (0, \underline{\sigma}]$, two countries set equilibrium tax rates in the following way:*

- (i) For $t_m \leq t_1^*$, $t_1^m = t_1(t_m)$ and $t_2^m \in [0, t_m]$;
- (ii) For $t_m > t_1^*$, two countries' equilibrium taxes are $(t_1^m = t_1(t_m), t_2^m \in [0, t_m])$ when $R_1(t_1(t_m), t_m) > R_1^m(\tilde{t}_1, t_2^m)$, $(t_1^m = \tilde{t}_1, t_2^m \in [0, \hat{t}_2^m])$ when $R_1(t_1(t_m), t_m) < R_1^m(\tilde{t}_1, t_2^m) < \bar{R}_1$, and $(t_1^m = \tilde{t}_1, t_2^m \in [0, 1])$ when $R_1^m(\tilde{t}_1, t_2^m) \geq \bar{R}_1$; both $(t_1(t_m), t_2^m \in [0, t_m])$ and $(\tilde{t}_1, t_2^m \in [0, t_m])$ are equilibrium tax rates when $R_1(t_1(t_m), t_m) = R_1^m(\tilde{t}_1, t_2^m)$,

where $\tilde{t}_i := \max \left\{ 0, \frac{(\sigma - \sigma_i^m)t_m}{\sigma} \right\}$, $R_1(t_1(t_m), t_m) = \varphi_1(t_1(t_m)) - \frac{t_1(t_m)[t_1(t_m) - t_m]}{\delta}$,

$$R_1^m(\tilde{t}_1, t_2^m) = \begin{cases} \frac{(\alpha_1 - r)^2}{2(2 - t_m)} & \text{if } \sigma > \sigma_1^m \\ \frac{(\alpha_1 - r)^2}{2(2 - t_m)} - \frac{(\sigma_1^m - \sigma)^2(2 - t_m)t_m^2}{2(1 - t_m)^2} & \text{if } \sigma \leq \sigma_1^m \end{cases}, \text{ and } \hat{t}_2^m \in (t_m, \bar{t}_1).$$

Proof. Given country 2's tax rate, country 1 can choose $t_1 = t_m$ to earn a positive revenue, i.e., $R_1^m(t_m, t_2) > 0$. By the same reasoning as in the proof of Proposition 2, we have:

$$\pi_1^m(t_1^m, t_2^m) > 0 \text{ and } t_1^m < \frac{\alpha_1 - r}{\alpha_1 - \mu r}.$$

We first prove part (i). From Lemma 3(i), when $t_m \leq t_1^*$, undercutting the minimum is a strictly dominated strategy for country 1. So its equilibrium tax rate must meet $t_1^m \geq t_m$. In what follows, we analyze two cases.

Case (i): assume that $0 \leq t_2^m \leq t_m$. Then repeating Step 2 in the proof of Proposition 2(ii), we have $t_1^m = t_1(t_m)$. On the other hand, given $t_1^m = t_1(t_m)$, country 2's revenue function is:

$$R_2^m(t_2, t_1^m) = \begin{cases} \frac{t_m(t_1^m - t_m)}{\delta} & \text{if } 0 \leq t_2 < t_m \\ \frac{t_2(t_1^m - t_2)}{\delta} & \text{if } t_m \leq t_2 < t_1^m \\ 0 & \text{if } t_2 \geq t_1^m \end{cases} . \text{ Together with } \frac{t_1^m}{2} < \frac{\alpha_1 - r}{2(\alpha_1 - \mu r)} \leq \frac{\alpha_2 - r}{\alpha_2 - \mu r} < t_m ,$$

we can show that any tax rate choice on interval $[0, t_m]$ is country 2's best response.

Case (ii): assume that $t_2^m > t_m$. In this case, notice that: $\forall t_1 \in [t_m, \min\{\bar{t}_1, t_2^m\}]$,

$R_1^m(t_1, t_2^m) = \varphi_1(t_1)$ with $\varphi_1'(t_1) > 0$. This implies that $t_1^m > t_m$. Recall that the only source of revenues for country 2 is taxing the profit shifted from country 1, which implies that country 2 undercuts country 1, i.e., $t_2^m < t_1^m$. So country 2's equilibrium revenue is

$R_2^m(t_2^m, t_1^m) = \frac{t_2^m(t_1^m - t_2^m)}{\delta}$. Then by the same reasoning as in Step 3 in the proof of Lemma 1,

for a tax rate t_2' that is marginally below t_2^m , we get: $R_2^m(t_2', t_1^m) > R_2^m(t_2^m, t_1^m)$, which contradicts the definition of Nash equilibrium. So there is no Nash equilibrium in this case.

Combining Cases (i) and (ii) leads to part (i).

Now we prove part (ii). When $t_m > t_1^*$, we analyze the following three cases.

Case (i): assume that $t_2^m \leq t_m$. By the same reasoning as in Step 2 in the proof of Proposition 2(iii), it is straightforward to show: $t_1^m = t_1(t_m)$ when $R_1(t_1(t_m), t_m) > R_1^m(\tilde{t}_1, t_2^m)$, and $t_1^m = \tilde{t}_1$ when $R_1(t_1(t_m), t_m) < R_1^m(\tilde{t}_1, t_2^m)$. Country 1 is indifferent between choosing $t_1(t_m)$ and $\tilde{t}_1$ when $R_1(t_1(t_m), t_m) = R_1^m(\tilde{t}_1, t_2^m)$. On the other hand, for $t_1^m = t_1(t_m)$, country 2's revenue function is the same as in Case (i) in the proof of part (i). For $t_1^m = \tilde{t}_1$, country 2 cannot attract any paper profit from country 1 such that $R_2^m(t_2, t_1^m) = 0$, $\forall t_2 \in [0, 1]$. Hence, any tax choice on interval $[0, t_m]$ is country 2's best response to t_1^m .

Case (ii): assume that $t_2^m > t_m$ and $t_1^m > t_m$. By the same reasoning as in case (ii) in the proof of Proposition D.1(i), we can show that no Nash equilibrium exists in this case.

Case (iii): assume that $t_2^m > t_m$ and $t_1^m \leq t_m$. From Lemma 3, it must be that $t_1^m = \tilde{t}_1$. On the other hand, for $t_1 \in [t_m, 1]$, country 1's revenue function depends on the relationship between t_2^m and $\frac{\alpha_1 - r}{\alpha_1 - \mu r}$ in the following way:

$$\text{When } t_2^m \geq \frac{\alpha_1 - r}{\alpha_1 - \mu r}, \quad R_1^m(t_1, t_2^m) = \begin{cases} \varphi_1(t_1) & \text{if } t_m \leq t_1 < \frac{\alpha_1 - r}{\alpha_1 - \mu r} \\ 0 & \text{if } t_1 \geq \frac{\alpha_1 - r}{\alpha_1 - \mu r} \end{cases}; \quad \text{When } t_2^m < \frac{\alpha_1 - r}{\alpha_1 - \mu r},$$

$$R_1^m(t_1, t_2^m) = \begin{cases} \varphi_1(t_1) & \text{if } t_m \leq t_1 < t_2^m \\ R_1(t_1, t_2^m) & \text{if } t_2^m \leq t_1 < \hat{t}_1 \\ 0 & \text{if } t_1 \geq \hat{t}_1 \end{cases}, \quad (\text{D1})$$

where $\hat{t}_1 \in \left(t_2^m, \frac{\alpha_1 - r}{\alpha_1 - \mu r}\right)$ is the unique value of t_1 such that

$$\pi_1^m(\hat{t}_1, t_2^m) = \left[f_1(k_1) - \mu r k_1 - \frac{t_1 - t_2^m}{\delta} \right]_{t_1 = \hat{t}_1} = 0.$$

Given (D1), solving the maximization problem $\max_{t_1 \in [t_m, 1]} R_1^m(t_1, t_2^m)$ yields:

$$\arg \max_{t_1 \in [t_m, 1]} R_1^m(t_1, t_2^m) = \begin{cases} t_1(t_2^m) & \text{if } t_m < t_2^m \leq t_2^\# \\ t_2^m & \text{if } t_2^\# < t_2^m \leq \bar{t}_1 \\ \bar{t}_1 & \text{if } t_2^m > \bar{t}_1 \end{cases}, \quad (\text{D2})$$

where $t_2^\# \in (t_1^N, \bar{t}_1)$ is the unique value of t_2 satisfying $t_1(t_2^\#) = t_2^\#$.

Plugging (D2) into (D1) yields the value function:

$$V(t_2^m) := \max_{t_1 \in [t_m, 1]} R_1^m(t_1, t_2^m) = \begin{cases} R_1(t_1(t_2^m), t_2^m) & \text{if } t_m < t_2^m \leq t_2^\# \\ \varphi_1(t_2^m) & \text{if } t_2^\# < t_2^m \leq \bar{t}_1 \\ \bar{R}_1 & \text{if } t_2^m > \bar{t}_1 \end{cases}, \text{ which is continuous in } t_2^m \text{ for all}$$

$t_2^m \in [t_m, 1]$ and strictly increases with t_2^m for all $t_2^m \in (t_m, \bar{t}_1)$.

Notice that $(\tilde{t}_1, t_2^m)$ are Nash equilibrium tax rates if and only if $R_1^m(\tilde{t}_1, t_2^m) \geq V(t_2^m)$,

$$\text{where } R_1^m(\tilde{t}_1, t_2^m) = \begin{cases} \frac{(\alpha_1 - r)^2}{2(2 - t_m)} & \text{if } \sigma > \sigma_1^m \\ \frac{(\alpha_1 - r)^2}{2(2 - t_m)} - \frac{(\sigma_1^m - \sigma)^2(2 - t_m)t_m^2}{2(1 - t_m)^2} & \text{if } \sigma \leq \sigma_1^m \end{cases}. \text{ Then it is straightforward to}$$

show the following:

If $R_1(t_1(t_m), t_m) < R_1^m(\tilde{t}_1, t_2^m) < \bar{R}_1$, then there exists a threshold $\hat{t}_2^m$ such that $t_1^m = \tilde{t}_1$, $t_2^m \in (t_m, \hat{t}_2^m]$ are equilibrium tax rates, where $\hat{t}_2^m \in (t_m, \bar{t}_1)$ is the unique value of t_2^m satisfying $V(\hat{t}_2^m) = R_1^m(\tilde{t}_1, \hat{t}_2^m)$;

If $R_1^m(\tilde{t}_1, t_2^m) \geq \bar{R}_1$, then $t_1^m = \tilde{t}_1$, $t_2^m \in (t_m, 1]$ are equilibrium tax rates;

If $R_1^m(\tilde{t}_1, t_2^m) \leq R_1(t_1(t_m), t_m)$, then there is no Nash equilibrium.

Lastly, combining Cases (i), (ii) and (iii) leads to part (ii) of the proposition. **Q.E.D.**

Appendix 3.E. The extension: including labor

In the main text, we have analyzed a simplified version of the GMT where capital is the only (variable) input in the MNE's production process. Now we extend the base model to include labor input and provide a complete picture of the GMT. Specifically, the affiliate in country i employs capital k_i and labor l_i to produce a homogenous good according to a general Cobb-Douglas production technology $f_i(k_i, l_i) := k_i^\lambda l_i^\beta$, with $\lambda, \beta \in (0, 1)$ and $\lambda + \beta < 1$. Each affiliate pays a wage rate w_i for each unit of labor. Labor is immobile, and the labor endowment in each country is $\bar{l}_i$ with $\bar{l}_1 > \bar{l}_2 > 0$. With this setup, country 1 (country 2) is the large country (small country) in the sense that country 1 has a larger population. The wage rate is endogenously determined by the labor market clearing condition $l_i = \bar{l}_i$, which equates labor demand l_i and labor endowment $\bar{l}_i$. In line with most countries' corporate tax systems, we assume that payroll cost is fully tax deductible and that a fraction $\mu \in [0, 1)$ of capital cost can be deducted from the corporate tax base. The GloBE income

(i.e., taxable profit) of affiliate i is $\pi_i = f_i(k_i) - \mu rk_i - w_i l_i + (-1)^i g$. The other settings in the base model are kept unchanged.

Absent the GMT, the MNE's after-tax profit is:

$$\Pi = \sum_{i=1}^2 \left[(1-t_i) \left(f_i - \mu rk_i - w_i l_i + (-1)^i g \right) - (1-\mu) rk_i \right] - \frac{\delta}{2} g^2.$$

The tax revenue of country i reads:

$$R_i = t_i \pi_i = t_i \left(f_i - \mu rk_i - w_i l_i + (-1)^i g \right).$$

Similar to the base model, we can establish the existence and uniqueness of the Nash equilibrium with $t_i^N \in \left(0, \frac{1-\lambda}{1-\mu\lambda} \right)$, $i = 1, 2$. At equilibrium, the small country sets a lower tax rate (i.e., $t_2^N < t_1^N$) to attract paper profits from the large country.

Now consider the introduction of the GMT with rate $t_m \in (t_2^N, t_1^N)$. In the short run, countries' tax rates are fixed, and only the MNE can adjust its strategies. The GMT is inactive for affiliate 1, while affiliate 2 has to pay a top-up tax, since country 2's tax rate is below the minimum. The SBIE allows the low-tax affiliate to deduct a fraction of the carrying value of tangible assets and payroll expenses from its GloBE income. For simplicity, we assume a common carve-out rate σ for capital stock and wage cost. Then the top-up tax paid by affiliate 2 is $(t_m - t_i) [\pi_i - \sigma(k_i + w_i l_i)]$.

The MNE's total after-tax profit reads:

$$\begin{aligned} \Pi^m &= \sum_{i=1}^2 \left[(1-t_i^N) \left(f_i - \mu rk_i - w_i l_i + (-1)^i g \right) - (1-\mu) rk_i \right] - (t_m - t_2^N) [f_2 - \mu rk_2 - w_2 l_2 + g - \sigma(k_2 + w_2 l_2)] - \frac{\delta}{2} g^2 \\ &= (1-t_1^N) (f_1 - \mu rk_1 - w_1 l_1 - g) - (1-\mu) rk_1 + (1-t_m) (f_2 - \mu rk_2 - w_2 l_2 + g) - (1-\mu) rk_2 \\ &\quad + \underbrace{\sigma(t_m - t_2^N)(k_2 + w_2 l_2)}_{\text{tax saved from the SBIE}} - \frac{\delta}{2} g^2. \end{aligned}$$

Under the QDMTT, the short-run revenues of the two countries are:

$$R_1^m = t_1^N (f_1 - \mu rk_1 - w_1 l_1 - g),$$

$$\begin{aligned} R_2^m &= t_2^N (f_2 - \mu rk_2 - w_2 l_2 + g) + (t_m - t_2^N) [f_2 - \mu rk_2 - w_2 l_2 + g - \sigma(k_2 + w_2 l_2)] \\ &= t_m (f_2 - \mu rk_2 - w_2 l_2 + g) - \underbrace{\sigma(t_m - t_2^N)(k_2 + w_2 l_2)}_{\text{loss from the deduction of SBIE}}. \end{aligned}$$

The short-run revenue effect of the GMT is given by the following:

Proposition E.1. *In the short run where countries' tax rates are fixed,*

- (i) *the large country benefits from the GMT;*
- (ii) *introducing a GMT with minimum rate marginally higher than the small country's equilibrium tax without the GMT increases (reduces) the small country's tax*

$$\text{revenue if } \phi_2(t_2^N) := \left[-\frac{\partial k_2 / k_2}{\partial t_2 / t_2} - \frac{t_2}{1-t_2} \frac{-\bar{l}_2 f_{2lk}''}{k_2 f_{2kk}''} - \frac{1}{1-t_2} \frac{w_2 \bar{l}_2}{k_2} - 1 \right] \Bigg|_{t_2=t_2^N} > 0 \quad (< 0).$$

This proposition holds for a general form of production function. $\phi_2(t_2^N)$ incorporates several key variables evaluated at country 2's tax rate t_2^N : the tax elasticity of capital investment (in absolute value) $-\frac{\partial k_2 / k_2}{\partial t_2 / t_2}$, the elasticity of substitution between labor and capital (in absolute value) $-\frac{\bar{l}_2 f_{2lk}''}{k_2 f_{2kk}''}$, and the ratio of payroll cost to capital stock $\frac{w_2 \bar{l}_2}{k_2}$.

Notably, Proposition E.1(ii) also applies to the base model where labor input is absent. Letting $\bar{l}_2 = 0$, we have $\phi_2(t_2^N) = -\frac{\partial k_2^N / k_2^N}{\partial t_2^N / t_2^N} - 1$. So when capital is the only (variable) production input, a marginal GMT reform increases (reduces) the short-run revenue of country 2 if the elasticity of capital investment in absolute value is greater (smaller) than unity. This is an equivalent statement of Proposition 1(ii) in the base model.

Given the Cobb-Douglas production technology, we can present that: $\forall i=1,2$,

$$\phi_i(t) = \left[-\frac{\partial k_i / k_i}{\partial t_i / t_i} - \frac{t_i}{1-t_i} \frac{-\bar{l}_i f_{ilk}''}{k_i f_{ikk}''} - \frac{1}{1-t_i} \frac{w_i \bar{l}_i}{k_i} - 1 \right] \Bigg|_{t_i=t} = \frac{t(1-\mu-\beta(1-\mu t))}{(1-\lambda)(1-t)(1-\mu t)} - \frac{\beta r(1-\mu t)}{\lambda(1-t)^2} - 1 =: \phi(t).$$

In the long run, both the MNE and the governments can adjust their behavior in response to the GMT. The following proposition characterizes the long-run equilibrium tax rates.

Proposition E.2. *After the introduction of the GMT, two countries set equilibrium tax rates in the following way:*

- (i) *For $\phi(t_m) \leq 0$, $t_1^m = t_1(t_m)$, $t_2^m = t_m$;*

- (ii) For $\phi(t_m) > 0$, two countries' equilibrium taxes are $(t_1^m = t_1(t_m), t_2^m = \tilde{t}_2)$ when $R_1(t_1(t_m), t_m) > R_1^m(\tilde{t}_1, \tilde{t}_2)$, and $(t_1^m = \tilde{t}_1, t_2^m = \tilde{t}_2)$ when $R_1(t_1(t_m), t_m) < R_1^m(\tilde{t}_1, \tilde{t}_2)$; both $(t_1(t_m), \tilde{t}_2)$ and $(\tilde{t}_1, \tilde{t}_2)$ are equilibrium tax rates when $R_1(t_1(t_m), t_m) = R_1^m(\tilde{t}_1, \tilde{t}_2)$, where $\tilde{t}_2 < t_m$.

When $\phi(t_m) \leq 0$, the GMT binds the small country and makes the large country choose tax rate along the initial best-response function. By contrast, when $\phi(t_m) > 0$, the small country will undercut the GMT rate and collect top-up taxes at equilibrium. Consider introducing a GMT with minimum rate marginally higher than the small country's equilibrium tax without the GMT. Then using Propositions E.1(ii) and E.2(i), we can conclude that: the marginal reform will raise the small country's revenue in the long run if it harms the small country in the short run, as in the base model (see Remark 5).

General Conclusion: Lessons for International Taxation of MNEs

This thesis has studied the effects of several developments of anti-avoidance policies on MNEs' behavior and countries' corporate taxes. The tax issues I focus on include coordinating tax enforcements, limiting tax deductibility and the OECD's global minimum tax. Based on the results of the thesis, I draw five lessons for the international taxation of multinationals.

First, the enforcement productivity (captured by the enforcement elasticity of tax revenue) plays a key role in coordinating tax enforcements, and low-enforcement productivity countries should move first in setting enforcements. On the other hand, a lack of local capacity in tax administration (e.g., designing tax legislation, administering tax laws, detecting revenue leakages) is a common feature of developing countries (see Albertin et al., 2021), which implies that low-enforcement productivity countries usually correspond to developing countries. Thus, our finding suggests that developing countries should take the lead in reinforcing anti-avoidance rules and make more efforts to combat profit shifting. The reason is that the amount of tax revenues at stake are larger for developing countries than for advanced economies if enforcement coordination fails.

Second, developing countries can benefit more from international enforcement cooperation than advanced economies, in the sense that the welfare gain relative to the non-cooperative outcome is larger for developing countries. Indeed, as argued by the OECD's (2015) Final Reports, "In a globalised economy, governments need to cooperate and refrain from harmful tax practices, to address tax avoidance effectively (...). Failure to achieve such cooperation (...) would have a disproportionately harmful impact on developing countries." Hence, it is of high importance for developing countries to further deepen cooperation and focus on monitoring the implementation and effectiveness of anti-avoidance measures. In particular, the multilateral instrument (Action 15) in the OECD/G20 BEPS project provides an innovative mechanism for governments to improve international tax cooperation.

Third, contingent on countries' tax capacity and industry-specific production technologies, tax deductibility affects countries' tax revenues differently. Under very low tax capacity, restricting the deductibility can lead to larger tax revenue. In contrast, under high tax capacity, low marginal production cost or large output elasticity, countries' equilibrium taxes strictly increase with tax deductibility. In practice, countries should maintain a balance between curbing profit shifting and improving production efficiency to effectively tax MNEs. The international tax design may differ across industries. 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 sector, the gain in revenue efficiency under turnover taxation is likely to be dominated by the great loss of production efficiency.

Fourth, introducing the GMT will reshape the international tax competition game and may not bring all countries' tax rates above the 15% rate. The SBIE reduces the top-up tax liability of affiliates with genuine economic activities and allows countries to provide tax incentives for these affiliates. Accordingly, small countries may still set their tax rates below the minimum rate, even at zero (if the carve-out rate is very small), in order to attract inward investments. Indeed, Ireland has decided to keep the corporate tax rate of 12.5% and top up the rate to 15% for Irish affiliates of foreign MNEs.

Fifth, the GMT can contain international profit shifting and raise large countries' tax revenues. However, under low costs of profit shifting (which might be the current situation due to cross-border tax loopholes), introducing the GMT will cause a short period of revenue losses for small countries. But they can benefit from a moderate GMT rate in the long run where countries' corporate tax rates are adjusted. This implies that it has to take some time until the benefits of the GMT realize for all countries.

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.
- Almunia, M., Lopez-Rodriguez, D., 2018. Under the radar: the effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy* 10(1), 1-38.
- Álvarez-Martínez, M. T., Barrios, S., d'Andria, D., Gesualdo, M., Nicodème, G., Pycroft, J., 2022. How large is the corporate tax base erosion and profit shifting? A general equilibrium approach. *Economic Systems Research* 34(2), 167-198.
- Amir, R., Stepanova, A., 2006. Second-mover advantage and price leadership in Bertrand duopoly. *Games and Economic Behavior* 55(1), 1-20.
- Auerbach, A. J., Devereux, M. P., 2018. Cash-flow taxes in an international setting. *American Economic Journal: Economic Policy* 10(3), 69-94.
- Baraké, M., Chouc, P. E., Neef, T., Zucman, G., 2022. Revenue effects of the global minimum tax under Pillar Two. *Intertax* 50(10), 689-710.
- Bauer, C. J., Langenmayr, D., 2013. Sorting into outsourcing: are profits taxed at a gorilla's arm's length? *Journal of International Economics* 90(2), 326-336.
- Becker, J., Davies, R. B., 2014. A negotiation-based model of tax-induced transfer pricing. *CESifo Working Paper* 4892, Munich.
- Beer, S., Loeprick, J., 2015. Profit shifting: drivers of transfer (mis)pricing and the potential of countermeasures. *International Tax and Public Finance* 22(3), 426-451.
- 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.
- Bucovetsky, S., 1991. Asymmetric tax competition. *Journal of Urban Economics* 30(2), 167-181.
- Bucovetsky, S., Haufler, A., 2008. Tax competition when firms choose their organizational form: should tax loopholes for multinationals be closed? *Journal of International Economics* 74(1), 188-201.
- Buettner, T., Poehnlein, M., 2024. Tax competition effects of a minimum tax rate: empirical evidence from German municipalities. *Journal of Public Economics* 236, 105148.
- Buettner, T., Wamser, G., 2013. Internal debt and multinational profit shifting: empirical evidence from firm-level panel data. *National Tax Journal* 66(1), 63-96.
- Bunn, D., Asen, E., Enache, C., 2020. *Digital Taxation around the World*. Washington, DC: Tax Foundation.
- Bustos, S., Pomeranz, D., Vila-Belda, J., Zucman, G., 2019. Challenges of monitoring tax compliance by multinational firms: evidence from Chile. *AEA Papers and Proceedings* 109, 500-505.
- Chen, X., Hindriks, J., 2023. Multinational taxation under pressure: the role of tax deductibility. *CORE/LIDAM Discussion Paper* 2023/13.
- 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.
- Cristea, A. D., Nguyen, D. X., 2016. Transfer pricing by multinational firms: new evidence from foreign firm ownerships. *The American Economic Journal: Economic Policy* 8(3), 170-202.
- 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.
- Department of Finance, 2023. Minister McGrath notes Ireland's application of effective 15% corporation tax rate for in-scope businesses. 31 December, 2023. <https://www.gov.ie/en/press-release/d4fcf-minister-mcgrath-notes-irelands-application-of-effective-15-corporation-tax-rate-for-in-scope-businesses-from-today/>.
- Devereux, M. P., Lockwood, B., Redoano, M., 2008. Do countries compete over corporate tax rates? *Journal of Public Economics* 92(5-6), 1210-1235.
- Devereux, M. P., Simmler, M., Vella, J., Wardell-Burrus, H., 2021. What is the substance-based carve-out under Pillar 2? And how will it affect tax competition? *EconPol Policy Brief* 39. ifo Institute, Munich.
- Devereux, M. P., Vella, J., Wardell-Burrus, H., 2022. Pillar 2: rule order, incentives, and tax

- competition. *Policy Brief*. Oxford University Centre for Business Taxation, Oxford.
- Dharmapala D., 2019. Profit shifting in a globalized world. *AEA Papers and Proceedings* 109, 488-492.
- 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.
- Durán-Cabré, J. M., Esteller-Moré, A., Salvadori, L., 2015. Empirical evidence on horizontal competition in tax enforcement. *International Tax and Public Finance* 22(5), 834-860.
- 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.
- Ernst & Young, 2015. Experiences with cash-flow taxation and prospects. *European Commission Taxation Papers, Working Paper* 55–2015.
- 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.
- European Commission, 2021. *Proposal for a Council Directive on Ensuring a Global Minimum Level of Taxation for Multinational Groups in the Union*, COM(2021) 823 final.
- European Commission, 2023. *Proposal for a Council Directive on Business in Europe: Framework for Income Taxation (BEFIT)*, COM(2023) 532 final.
- Faulkender, M. W., Hankins, K. W., Petersen, M. A., 2019. Understanding the rise in corporate cash: Precautionary savings or foreign taxes. *The Review of Financial Studies* 32(9), 3299-3334.
- Ferrari, A., Laffitte, S., Parenti, M., Toubal, F., 2023. Profit shifting frictions and the geography of multinational activity. *CEPR Discussion Papers* (17801).
- Gaspar Committee, 2014. *Report of the Commission Expert Group on Taxation of the Digital Economy*, chaired by Vitor Gaspar, Brussels: European Commission.
- Gresik, T. A., 2016. Allowing firms to choose between separate accounting and formula apportionment taxation. *Journal of Public Economics* 138, 32-42.
- Gumpert, A., Hines Jr, J. R., Schnitzer, M., 2016. Multinational firms and tax havens. *Review of Economics and Statistics* 98(4), 713-727.
- Hamilton, J. H., Slutsky, S. M., 1990. Endogenous timing in duopoly games: Stackelberg or Cournot equilibria. *Games and Economic Behavior* 2(1), 29-46.

- Harsanyi, J.C., Selten, R., 1988. *A General Theory of Equilibrium Selection in Games*. MIT Press, Cambridge, MA.
- Haufler, A., Kato, H., 2024. A global minimum tax for large firms only: implications for tax competition. *CESifo Working Paper* No. 11087.
- Haufler, A., Mardan, M., Schindler, D., 2018. Double tax discrimination to attract FDI and fight profit shifting: The role of CFC rules. *Journal of International Economics* 114, 25-43.
- Hebous, S., Keen, M., 2023. Pareto-improving minimum corporate taxation. *Journal of Public Economics* 225, 104952.
- Hindriks, J., Nishimura, Y., 2015. A note on equilibrium leadership in tax competition models. *Journal of Public Economics* 121, 66-68.
- Hindriks, J., Nishimura, Y., 2017. Equilibrium leadership in tax competition models with capital ownership: a rejoinder. *International Tax and Public Finance* 24(2), 338-349.
- Hindriks, J., Nishimura, Y., 2021. Taxing multinationals: the scope for enforcement cooperation. *Journal of Public Economic Theory* 23(3), 487-509.
- 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.
- Hugger, F., Cabral, A. C. G., Bucci, M., Gesualdo, M., O'Reilly, P., 2024. The global minimum tax and the taxation of MNE profit. *OECD Taxation Working Papers* No. 68.
- Huizinga, H., Laeven, L., 2008. International profit shifting within multinationals: a multi-country perspective. *Journal of Public Economics* 92(5-6), 1164-1182.
- IMF, 2023. International corporate tax reform. *IMF Policy Paper* No. 2023/001.
- Janeba, E., Schjelderup, G., 2023. The global minimum tax raises more revenues than you think, or much less. *Journal of International Economics* 145, 103837.
- Johannesen, N., 2010. Imperfect tax competition for profits, asymmetric equilibrium and beneficial tax havens. *Journal of International Economics* 81(2), 253-264.
- Johannesen, N., 2022. The global minimum tax. *Journal of Public Economics* 212, 104709.
- Kanbur, R., Keen, M., 1993. Jeux sans frontieres: tax competition and tax coordination when countries differ in size. *The American Economic Review* 83(4), 877-892.
- Keen, M., Konrad, K. A., 2013. The theory of international tax competition and coordination. *Handbook of Public Economics* 5, 257-328.
- 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., Slemrod, J., 2017. Optimal tax administration. *Journal of Public Economics* 152, 133-142.

- Keen, M., Wildasin, D., 2004. Pareto-efficient international taxation. *American Economic Review* 94(1), 259-275.
- Kempf, H., Rota-Graziosi, G., 2010. Endogenizing leadership in tax competition. *Journal of Public Economics* 94(9-10), 768-776.
- 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.
- Kleinbard, E. D., 2013. Through a latte, darkly: Starbucks's stateless income planning. *Tax Notes*, 24 June, 1515-1535.
- Koethenbuerger, M., Mardan, M., Stimmelmayer, M., 2019. Profit shifting and investment effects: the implications of zero-taxable profits. *Journal of Public Economics* 173, 96-112.
- 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., Stimmelmayer, M., 2018. Tax revenue losses through cross-border loss offset: an insurmountable hurdle for formula apportionment? *European Economic Review* 102, 188-210.
- Mardan, M., Stimmelmayer, M., 2020. Tax competition between developed, emerging, and developing countries – Same same but different? *Journal of Development Economics* 146, 102491.
- 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, 2015. *BEPS Project Explanatory Statement: 2015 Final Reports*. Paris: OECD Publishing.
- OECD, 2018. *Tax Challenges Arising from Digitalisation – Interim Report 2018*. Paris: OECD Publishing.
- OECD, 2021. *Tax Challenges Arising from the Digitalisation of the Economy – Global Anti-Base Erosion Model Rules (Pillar Two): Inclusive Framework on BEPS*. Paris: OECD Publishing.
- OECD, 2023. *The Multilateral Convention to Implement Amount A of Pillar One*. Paris: OECD Publishing.
- Overesch, M., Rincke, J., 2011. What drives corporate tax rates down? A reassessment of

- globalization, tax competition, and dynamic adjustment to shocks. *The Scandinavian Journal of Economics* 113(3), 579-602.
- Palmer, J. A., 2003. Relative convexity. *ECE Dept., UCSD, Tech. Rep.*
- Peralta, S., van Ypersele, T., 2005. Factor endowments and welfare levels in an asymmetric tax competition game. *Journal of Urban Economics* 57(2), 258-274.
- Peralta, S., Wauthy, X., Van Ypersele, T., 2006. Should countries control international profit shifting? *Journal of International Economics* 68(1), 24-37.
- Perry, V. J., 2023. Pillar 2: tax competition in low-income countries and substance-based income exclusion. *Fiscal Studies* 44(1), 23-36.
- Riedel, N., 2011. Taxing multi-nationals under union wage bargaining. *International Tax and Public Finance* 18(4), 399-421.
- Riedel, N., Runkel, M., 2007. Company tax reform with a water's edge. *Journal of Public Economics* 91(7-8), 1533-1554.
- Riedel, N., Zinn, T., Hofmann, P., 2015. Do transfer pricing laws limit international income shifting? Evidence from Europe. *CESifo Working Paper* 4404, Munich.
- Schjelderup, G., Stähler, F., 2023. The economics of the global minimum tax. *International Tax and Public Finance* 1-18.
- Slemrod, J., Wilson, J. D., 2009. Tax competition with parasitic tax havens. *Journal of Public Economics* 93(11-12), 1261-1270.
- Slemrod, J., Yitzhaki, S., 2002. Tax avoidance, evasion, and administration. *Handbook of Public Economics* 3, 1423-1470.
- 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., 2023. The missing profits of nations. *The Review of Economic Studies* 90(3), 1499-1534.
- UNCTAD, 2022. International tax reforms and sustainable investment. *World Investment Report* 2022.
- Van Huyck, J. B., Battalio, R. C., Beil, R. O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review* 80(1), 234-248.
- Vives, X., 2000. *Oligopoly Pricing: Old Ideas and New Tools*. MIT Press, Cambridge, Massachusetts.
- Wang, Y. Q., 1999. Commodity taxes under fiscal competition: Stackelberg equilibrium and optimality. *American Economic Review* 89(4), 974-981.
- Wier, L. S., Zucman, G., 2022. Global profit shifting, 1975-2019. *NBER Working Paper* No. 30673.

- Zinn, T., Riedel, N., Spengel, C., 2014. The increasing importance of transfer pricing regulations: a worldwide overview. *Intertax* 42(6/7), 352-404.
- Zucman, G., 2014. Taxing across borders: tracking personal wealth and corporate profits. *Journal of Economic Perspectives* 28(4), 121-148.