University of Toronto
Department of Economics

Working Paper 731

A Macroeconomic Perspective on Taxing Multinational Enterprises

By Sebastian Dyrda, Guangbin Hong and Joseph B Steinberg

September 01, 2022

A Macroeconomic Perspective on Taxing Multinational Enterprises*

Sebastian Dyrda
University of Toronto

Guangbin Hong
University of Toronto

Joseph B. Steinberg
University of Toronto

September 1, 2022

Abstract

We develop a framework to study the macroeconomic implications of taxing multinational enterprises (MNEs) that shift profits to subsidiaries in low-tax jurisdictions by transferring ownership of non-rival intangible capital. We first prove analytically that profit shifting increases intangible investment, leading to higher profits and output at the MNE level. We then calibrate our model so that it reproduces salient country-level facts about production, trade, FDI, and, most importantly, profit shifting. We use our calibrated model to evaluate the consequences of two proposals by the OECD and G20 governments to reduce profit shifting by MNEs: allocating the rights to tax some of an MNE's profits to the countries in which it sells its products; and a 15% minimum global corporate income tax. We show that these policies would reduce profit shifting by more than two-thirds, but would also reduce intangible investment and output in high-tax regions. This highlights a key tension for policymakers: profit shifting erodes high-tax countries' tax bases, but also boosts economic activity, and thus policies that reduce profit shifting have harmful macroeconomic side effects.

Keywords: Multinational enterprise; transfer pricing; profit shifting; base erosion; intangible capital; corporate tax.

JEL Codes: E6, F23, H25, H27

*Steinberg: joseph.steinberg@utoronto.ca; Hong: g.hong@mail.utoronto.ca; Dyrda: sebastian.dyrda@utoronto.ca. Address: University of Toronto, Max Gluskin House, 150 St. George Street, Toronto, ON M5S 3G7, Canada. Acknowledgments: We thank Ellen McGrattan, V.V. Chari, Burhan Kuruscu, Diego Restuccia, Florian Sheuer, and Aleh Tsyvinski for useful comments and discussions. We also thank participants of the NBER Summer Institute Macro Public Finance workshop, Society for Economic Dynamics 2022 Annual Meeting, and Midwest Macro Spring 2022 conference.

1 Introduction

In October 2021, 137 countries agreed to implement the largest international tax reform in history to address the growing problem of profit shifting by multinational enterprises (MNEs). We show that regulating profit shifting would have a significant economic cost: it would reduce MNEs' investment in intangible capital, and in equilibrium this would cause aggregate output to fall. Using a new theory that links profit shifting by MNEs to their intangible investment decisions, we demonstrate the real implications of this phenomenon and quantify the adverse macroeconomic impact of the landmark policy designed to eliminate it.

Base erosion and profit shifting (BEPS), as it is called by the OECD, refers to MNEs' use of tax planning strategies to exploit gaps and mismatches in tax rules to artificially shift profits to low- or no-tax countries where they conduct little or no economic activity, or to erode tax bases through deductible payments such as interest or royalties. The scale of profit shifting is striking. For example, Tørsløv, Wier, and Zucman (2022) estimate that 36 percent of worldwide multinational profits are shifted to tax havens, while Guvenen, Mataloni, Rassier, and Ruhl (2022) find that 38 percent of foreign income reported by U.S. MNEs is actually generated at home in the United States. The implications for public finances are equally striking: Clausing (2020a) estimates that about a third of U.S. corporate income taxes are lost to profit shifting, which is equivalent to more than $100 billion per year. According to the OECD, it reduces global corporate income tax revenues by as much as 10 percent per year, or $240 billion (Johansson et al., 2017). Consequently, addressing this issue is a top priority for policymakers in high-tax countries where many of the biggest MNEs are based. The OECD/G20 Inclusive Framework on BEPS outlines two major policy changes, or "pillars".[1] The first pillar is revenue-based profit allocation, which allocates the rights to tax some of an MNE's profits to the countries in which it operates in proportion to these countries' shares of the MNE's global sales. The second is a global minimum corporate income tax, which would require that all corporate income, regardless of where it is booked, be taxed at no lower than 15 percent.

In order to study the macroeconomic effects of profit shifting and the two-pillar OECD policy designed to address it, we develop a model in which MNEs shift profits by transferring intangible capital property rights to foreign subsidiaries in low-tax countries.[2] As in McGrattan and Prescott (2010), intangible capital is non-rival: MNEs produce it by doing

[1]The press statement and a description of these pillars can be found here.

[2]See Guvenen et al. (2022) and Delis et al. (2021) for evidence on the role of intangible capital in profit shifting.

research and development at home, but use it to produce simultaneously in all of their foreign subsidiaries around the world. According to transfer pricing rules, these subsidiaries pay licensing fees to use this capital. Normally, these fees are paid to the domestic parent corporation, but the rights to this capital can be transferred—at a cost—to subsidiaries in low-tax regions, i.e., tax havens. The income that accrues to the MNE's intangible capital is now taxed at a lower rate, increasing the after-tax return on intangible investment. This increases the MNE's optimal level of research and development, which ultimately leads the MNE to produce more output both at home and abroad.

We use our model to make two substantive contributions, one theoretical and one quantitative. In the theoretical part of the paper, we characterize analytically an impact of profit shifting on MNEs' production decisions and profitability. We prove that profit shifting increases intangible investment, leading to higher output but lower reported profits in the MNE's home country. This result clearly reveals the tradeoff that profit shifting presents to policymakers: although it artificially redistributes MNEs' income to foreign tax havens, it also increases the amount of income that they actually generate. Moreover, the size of this effect is increasing in the difference between the corporate tax rate in the MNE's home country and the tax haven's tax rate; increasing the tax rate in the tax haven reduced intangible investment at home. This has direct implications for the second pillar of the OECD/G20 plan: a global minimum tax rate will reduce intangible investment, and the higher the minimum tax rate, the larger the reduction. Further, we prove that sales-based profit reallocation, the first pillar of this plan, will also have adverse effects on real economic activities.

In the quantitative part of the paper, we embed our theory of profit shifting into a general-equilibrium environment to measure the macroeconomic effects of the OECD/G20 BEPS framework. Our quantitative model features five regions that differ in population, productivity, and corporate tax rates. We split the countries identified as tax havens by Tørsløv et al. (2022) into two regions: a productive low-tax region that includes Ireland, Switzerland, and other countries where most of the economy is not devoted to profit shifting; and a tax haven that includes the Caribbean, the Channel Islands, and other small countries whose economies rely heavily on profit shifting. The other three regions are North America, Europe (minus countries in the low-tax region), and the rest of the world. Heterogeneous firms pay fixed costs to establish foreign subsidiaries in other regions (i.e., become MNEs) as in Helpman, Melitz, and Yeaple (2004) and Garetto, Oldenski, and Ramondo (2019), and firms with subsidiaries in one of the first two regions can shift profits by transferring the rights to intangible capital according to our theory. We discipline our model using data on production, trade, multinational activity, and, most importantly, estimates of international profit shifting from Tørsløv et al. (2022). We find that the OECD's proposal would go a long way toward eliminating profit shifting: lost profits would fall by 77 percent in North America, 82 percent in Europe, and 90 percent in the rest of the world. However, it would also materially reduce intangible capital investment and overall macroeconomic performance across the globe: GDP would fall by 0.17 percent in North America, 0.16 percent in Europe, 0.13 percent in the low-tax productive region, and 0.14 percent in the rest of the world.

2 Related literature

This paper contributes to several strands of literature. First, there are a number of studies that aim to measure the scope of international profit shifting by MNEs. Guvenen, Mataloni, Rassier, and Ruhl (2022) use confidential survey data from the Bureau of Economic Analysis (BEA) and estimate that 38 percent of income recorded by U.S. MNEs on their foreign direct investment should be re-attributed to U.S. GDP. Importantly, they document that profit shifting is concentrated in industries and firms with significant research and development spending, providing support for our theory that intangible assets are central to profit shifting. Tørsløv, Wier, and Zucman (2022) combine cross-country data on wages and profitability of foreign firms' local affiliates (a.k.a. foreign affiliates statistics) versus local firms. Their main finding is that 36% of multinational profits were shifted to tax havens globally in 2015. Clausing (2020a) conclude, based on several different estimates, that the U.S. tax revenue loss from profit shifting in 2017 likely exceeded $100 billion, or about a third of federal corporate tax revenues.[3] Our interpretation of these estimates is that profit shifting is a large and consequential problem at the global scale.[4] Our contribution to this strand of literature is to develop a quantitative macroeconomic framework to assess the impact of transfer pricing and profit shifting on macroeconomic aggregates and tax revenues. We exploit the empirical estimates discussed above to discipline our structural model.

We also contribute to the literature on the macroeconomic role of intangible capital. The importance of the intangible capital for aggregate measurement has been highlighted by Corrado, Hulten, and Sichel (2009), McGrattan and Prescott (2010), O'Mahony, Corrado, Haskel, Iommi, Jona-Lasinio, and Mas (2018), and Koh, Santaeulàlia-Llopis, and Zheng

[3]These findings are in line with the $4.2 trillion in offshore earnings observed in the U.S. data, $3.0 trillion of which is in known tax havens or countries with effective tax rates below 10 percent.

[4]Blouin and Robinson (2020) and Clausing (2020a) discuss the methodological challenges associated with estimating the magnitude of profit shifting. Bolwijn, Casella, and Rigo (2018) and Crivelli, De Mooij, and Keen (2015) study the impact of profit shifting on tax revenues for developing countries. See Dowd, Landefeld, and Moore (2017), Clausing (2016), and OECD (2015) for extensive reviews of the profit-shifting literature and the estimates found therein.

(2020); Peters and Taylor (2017) and Ewens, Peters, and Wang (2019) demonstrate its importance for firm-level measurement. Importantly, Delis, Delis, Laeven, and Ongena (2021) establish a causal, positive relationship at the firm level between profit shifting and the share total assets that are intangible.[5] On the modelling front, McGrattan and Prescott (2010) develop a multi-country general equilibrium model that includes two types of intangible capital: rival, plant-specific intangible capital; and non-rival technology capital that can be used in multiple locations simultaneously. They use their model to explain the differences between the investment returns of foreign subsidiaries of U.S. multinational companies and the returns of U.S. subsidiaries of foreign multinationals. We contribute to this line of research by developing a novel model of transfer pricing and profit shifting that centers around nonrival intangible capital. In our framework, a positive relationship between the ratio of intangible capital to total assets and profit shifting arises endogenously from MNEs' decisions.

This paper is also related to the macro public finance literature on corporate income taxation. This strand of research is vast and dates back to seminal contributions by Harberger (1962) and Auerbach (1983), among others. More recently, Barro and Furman (2018) assess the macroeconomic consequences of the Tax Cuts and Jobs Act of 2017 (TCJA). Kaymak and Schott (2018) argue that falling corporate income taxes across the world are the main driver behind the decline of the labor share. Kaymak and Schott (2019) argue that loss-offset provisions in the corporate income tax code give rise to capital misallocation and assess the associated aggregate output losses. Bhandari and McGrattan (2020) quantify the impact of reducing corporate income taxes in a model where firms choose the legal form of business organization endogenously. In spite of the large number of studies in this literature, little attention has been paid to the macroeconomic effects of international profit shifting and its impact on intangible investment. This paper aims to fill this gap.

Furthermore, our work contributes to the the international economics literature on multinational firms. See Antrás and Yeaple (2014) for a review of this extensive line of research. We extend the Helpman et al. (2004) framework by incorporating both non-rival, intangible capital and profit shifting into the multinational firm's decision problem.[6] Both features are central for understanding the growing impact of multinational firms on the global economy.

Lastly, our paper relates to a large, influential literature on international tax competition.[7]

[5] Specifically, they estimate that a one-standard-deviation increase in the ratio of intangible assets to total assets increases profit shifting by approximately 3 percent. The intangible ratio is the characteristic with the largest impact on profit shifting in their analysis.

[6] Our model also shares some ingredients with work by Melitz (2003), Chaney (2008), Garetto, Oldenski, and Ramondo (2019), and McGrattan and Waddle (2020).

[7] The notion of international tax competition originates from a theoretical literature on capital tax competition across jurisdictions, which has roots back to Tiebout (1956) but took shape with the seminal papers

The growing importance of profit shifting has led to a rapid development of this literature in recent years; see Keen and Konrad (2013) for a review. Among the most important papers in this line of research are Haufler and Schjelderup (2000), Mintz and Smart (2004), Hong and Smart (2010), Slemrod and Wilson (2009) and Johannesen (2022). In this paper, we do not follow the game-theoretic approach usually pursued in this literature, but we explicitly incorporate an endogenous profit-shifting margin into the multinational firm's problem.[8]

3 Institutional Background

In this section we provide a brief overview of the current international tax regime and describe the main features of the two-pillar reform proposed by the OECD. We aim here to deliver an executive summary, rather than an exhaustive discussion, of these immensely complex issues.[9] Understanding the main components of the international tax architecture is crucial since they largely dictate the setup of our theory and impose restrictions on what any reform proposal can achieve.

3.1 The Current International Tax Regime

The existing international law entitles the country to tax persons, either natural or legal, with which it has sufficient ties. In practice the taxing rights are a product of multiple national laws and international treaties often contradicting one another. The following are the most important characteristics of the current regime.

Legal separation of entities. The current regime treats subsidiaries within one MNE as separate legal entities. Thus, any transaction between parts of an MNE in different tax jurisdictions, such as for example an asset purchase, has real tax consequences. This characteristic coupled with heterogeneity of the tax systems across jurisdictions and transfer prices' manipulation gives rise to profit shifting opportunities.

Allocation of taxing rights. There are at least four possible locations where multinational companies might in principle be taxed: the location of its shareholders, parent companies, affiliates, or customers. According to the current regime MNEs are taxed primarily in the third location (affiliates' location), but sometimes also in the second. This is achieved by a

of Zodrow and Mieszkowski (1986) and Wildasin (1988).

[8] In a separate paper, we incorporate our theory of profit shifting into the standard tax competition framework. See Dyrda, Hong, and Steinberg (2022).

[9] Our summary is largely based on: Devereux et al. (2021), OECD (2015), OECD (2017) and OECD (2022).

combination of legal rules allowing the countries to tax according to the source or residence basis.[10]

Transfer prices. The within-MNE transactions occur at transfer prices, which are disciplined by the so-called *arm's length principle (ALP)*. The basic idea behind the ALP is that within-MNE prices should reflect the market prices that would have been charged by two independent parties of transactions. There are five core methods to achieve the ALP standard: the comparable uncontrolled price (CUP), resale price minus, cost plus, profit split, and transactional net margin method. The practical implementation of this principle is challenging and requires complex guidelines published regularly by the OECD which member countries should obey–see OECD (2022) for the latest guide.

Treatment of intangibles. The method preferred by the OECD to implement ALP is CUP, which simply employs the price charged on comparable transactions between independent parties. CUP however is hard to implement in case of trading intangibles, since most of the time a comparable transaction is non-existent. In such cases the preferred method is the profit split method, which in short inspect the relative financial or other contributions made by the two companies entering into a transaction. A profit split is then determined based on these contributions. OECD (2014) provides extensive guidelines on pricing transactions involving intangibles.

3.2 OECD Base Erosion and Profit Shifting Project

In what follows we briefly summarize the key provisions of reform proposed by OECD/G20 Inclusive Framework on BEPS, as they were at the time of the writing of this paper.[11]

3.2.1 Pillar One: Profit allocation and nexus.

The general principle behind the Pillar One is to allocate taxing rights more closely with where the customers and users of the in-scope MNEs are located. The key elements of the Pillar One are as follows.

[10] From a legal perspective, a country taxes on a residence basis when it taxes companies that are resident in that country for tax purposes on income arising in that or in another country. A country taxes on a source basis when it taxes companies that are not resident in that country for tax purposes on income deemed to arise in that country. For a thorough discussion of these concepts see Devereux et al. (2021).

[11] The details of both pillars as well as the exact implementation plan has been very much "work in progress" at the time of writing this paper. Since November 2021 OECD has been organizing a series of public consultation meetings in order to work out technical details and parameters of the reform.

Scope. The new profit allocation rule will apply to groups with greater than €20 billion in worldwide revenues and a profitability before tax margin of at least 10 percent. There are some exclusions for extractive industry and regulated financial services.

Nexus. The allocation key is based on the revenue that is sourced to each jurisdiction. It will be sourced to the end market jurisdictions where goods or services are used or consumed. It permits allocation to a market jurisdiction where the in-scope MNE derives at least €1 million in revenues from that jurisdiction.

Quantum. For in-scope MNEs, 25% of residual profit, i.e. profit in excess of 10% of revenue, will be allocated to market jurisdictions with nexus using a revenue-based allocation key.

Elimination of double taxation. Profit allocated to a market jurisdiction will be dispensed from double taxation through direct exemption of credit method.

Unilateral Measures. The agreement requires all parties to remove all Digital Services Taxes and other relevant, similar measures with respect to all companies and to commit not to introduce such measures in the future.

3.2.2 Pillar Two: Global minimum taxation.

The second pillar consists of two sets of rules granting jurisdictions additional taxing rights: (i) interlocking domestic rules labelled as Global anti-Base Erosion Rules (GloBE) (ii) a treaty-based rule Subject to Tax Rule (STTR). Their key features are as follows.

Scope. GloBe rules apply to multinational enterprise groups with a total consolidated group revenue above €750 million in at least two of the four preceding years.

Minimum tax rate. GloBE rules apply a system of top-up taxes that brings the total amount of taxes paid on an MNE's profit in a jurisdiction up to the minimum rate of 15%.

Exclusions. GloBe rules will also provide for an exclusion for those jurisdictions where the MNE has revenues of less than EUR 10 million and profits of less than EUR 1 million.

Subject to Tax Rule (STTR). It complements the GloBE rules by targeting intra-MNE payments exploiting certain provisions of the treaty to shift profits from source countries to payee jurisdictions, where those payments are subject to no or low rates of nominal taxation. In such cases, it reallocates taxing rights to source jurisdictions. It applies to such payments as covered payments—interest, royalties, brokerage, marketing, procurement, agency or other intermediary services, etc. The minimum rate for the STTR will be 9 percent.

4 Theory of Profit Shifting and Intangible Investment

In order to study the real effects of international profit shifting and the OECD/G20 policy framework designed to address this phenomenon, we develop a theory that links profit shifting to intangible investment. According to our theory, MNEs shift profits by transferring the rights to non-rival intangible capital to subsidiaries in low-tax jurisdictions. This increases the after-tax return on intangible capital, which leads MNEs to increase their investment in this capital in equilibrium and ultimately produce more output. Thus, profit shifting presents policymakers in high-tax countries with a trade-off: it reduces corporate income tax revenues, but also increases overall economic activity.

4.1 Environment

Consider an MNE that operates subsidiaries in I regions. Each region $k = 1, \ldots, I$ is characterized by population N_k, total factor productivity A_k, and corporate tax rate $\tau_k \in [0, 1]$. The MNE's home region is denoted by i. Without loss of generality, we normalize the entire population across regions to unity, i.e. $\sum_{k=1}^{I} N_k = 1$. We refer to the region with the lowest tax rate, which we denote by i^*, as the tax haven, i.e., $\tau_{i^*} = \min\{\tau_1, ..., \tau_I\}$.

In each region, the MNE has access to a production technology F_k in that transforms labor ℓ_k and intangible capital z into a final good:

$$F_k(z, l_k) = A_k (N_k z)^{\phi} l_k^{\gamma}. \tag{1}$$

As in McGrattan and Prescott (2010), intangible capital is non-rival: it is purchased in the headquarters region i at the local price p_i, but it can be used in all I locations simultaneously. Its productivity is determined by the local population N_k, which proxies for the number of production locations in a given region where the intangible capital can be deployed. Labor is rented in a competitive market at wage rate w_k. We assume decreasing returns to scale, i.e., $\phi + \gamma < 1$.[12]

As a starting point, we begin by defining the MNE's profits in the standard setup (e.g., as in McGrattan and Prescott, 2010) in which foreign subsidiaries use intangible capital free

[12] In our quantitative model we assume constant returns to scale and monopolistic competition. In this partial equilibrium setting, the two approaches are isomorphic. We choose decreasing returns here for its analytical simplicity.

of charge:

$$\pi_i = p_i \left(A_i \left(N_i z \right)^{\phi} l_i^{\gamma} \right) - w_i l_i - p_i z \tag{2}$$

$$\pi_k = p_k \left(A_k \left(N_k z \right)^{\phi} l_k^{\gamma} \right) - w_k l_k, \quad \forall k \neq i. \tag{3}$$

We refer to this as the *free transfer* (FT) scenario and denote the allocation of intangible capital in this case by z^{FT}. Our methodological innovation is to add two new ingredients to this setup: transfer pricing and profit shifting, which we do one at a time.

In the *transfer pricing* (TP) scenario, the parent division retains legal ownership of the MNE's stock of intangible capital and licenses the right to use this capital to its foreign affiliates. The accounting profits in each of the MNE's divisions in this scenario are

$$\pi_i^{TP} = \pi_i + \sum_{k \neq i} \vartheta_k \left(z \right) z, \tag{4}$$

$$\pi_k^{TP} = \pi_k - \vartheta_k \left(z \right) z \quad \forall k \neq i. \tag{5}$$

According to the arm's length principle, the licensing fees, ϑ_k, are set to the affiliates' marginal revenue products of intangible capital, i.e.,

$$\vartheta_k \left(z \right) \equiv \phi p_k N_k \left(A_k \left(N_k z \right)^{\phi - 1} l_k^{\gamma} \right). \tag{6}$$

We denote the allocation of intangible capital in this case by z^{TP}. In this section, we assume that the MNE takes $q_k(z)$ as given according to the spirit of the arm's length principle; it does not internalize the effect of its choice of z on q_k. Mathematically speaking, the MNE does not take the derivative of q_k when taking the first-order condition of its profit function with respect to z. This keeps our key equations relatively simple, which allows us to highlight the important economic forces at work behind our results. In the appendix, we show that all of our analytical results hold when the MNE does internalize the effect of its choice of z on q_k, and we allow for this effect in our quantitative analysis as well.

In the *profit shifting* (PS) scenario, the MNE's headquarter sells a fraction λ of its intangible capital to its affiliate in the tax haven, which then licenses the rights to use this capital to the parent division and the other non-haven foreign affiliates. We assume that the tax-haven affiliate buys intangible capital from the headquarters at a markdown $\varphi \leq 1$ below the competitive price, which is equal to the sum total of the licensing fees that this capital can generate, i.e., the sum of the marginal revenue products across all of the regions in which the MNE operates. Manipulating transfer prices in this way is assumed to be costly, as the multinational needs to modify its books, and possibly its real trade and investment patterns, to be able to justify the distorted transfer prices to the tax authorities. We impose the following assumption on the cost function $\mathcal{C}(\lambda)$.

Assumption 1 *Let* $\mathcal{C}(\lambda) \equiv \lambda + (1-\lambda)\log(1-\lambda)$*, implying* $\mathcal{C}'(\lambda) = -\log(1-\lambda)$*,* $\mathcal{C}(0) = 0$*,* $\mathcal{C}(1) = 1$*, and* $\lambda \in [0,1]$*.*

It is important to note that $C(\lambda)$ captures direct costs of profit shifting (e.g. increased spending on lawyers, accountants, and transfer pricing consultants), but also, in a reduced-form way, the increased risk of penalization by the government (see, e.g., Allingham and Sandmo, 1972; Rotberg and Steinberg, 2022).

Pre-tax profits in the profit shifting scenario are thus:

$$\pi_i^{PS} = \pi_i + z\left[\varphi\lambda\sum_k \vartheta_k(z) - \lambda\vartheta_i(z) + (1-\lambda)\sum_{k\neq i}\vartheta_k(z) - \sum_k \vartheta_k(z)\,\mathcal{C}(\lambda)\right], \tag{7}$$

$$\pi_{i^*}^{PS} = \pi_{i^*} + z\left[\lambda\sum_{k\neq i^*}\vartheta_k(z) - (1-\lambda)\,\vartheta_{i^*}(z) - \varphi\lambda\sum_k \vartheta_k(z)\right], \tag{8}$$

$$\pi_k^{PS} = \pi_k - \vartheta_k(z)\,z \quad \forall k \neq i. \tag{9}$$

The first term in square brackets in (7), $\varphi\lambda\sum_k \vartheta_k(z)$, is the revenue from selling intangible capital to the tax haven. The second term, $-\lambda\vartheta_i(z)$, denotes the licensing fee that the headquarter pays to the tax haven for the right to use the fraction λ of intangible capital that has changed ownership. The third term, $(1-\lambda)\sum_{k\neq i}\vartheta_k(z)$, represents the licensing fees that the headquarter collects from the other affiliates for the remaining intangible capital that the headquarter retains. The term $\mathcal{C}(\lambda)\sum_k \vartheta_k(z)$ captures the costs of shifting intangible capital to the tax haven. The terms in (8) have analogous interpretations. We denote the allocation of intangible capital in this scenario by z^{PS}.

Consider the problem of maximizing after-tax profits in each scenario:

$$\max_{z^s, \left\{\ell_k^s\right\}_{k=1}^I, \lambda} \sum_{k=1}^I (1-\tau_k)\,\pi_k^s \tag{10}$$

where $s \in \{FT, TP, PS\}$. Note that λ is only chosen in the profit shifting scenario. We first characterize the MNE's optimal choice of λ in this scenario, and then characterize how this choice alters the MNE's intangible investment decision. The formal proofs of these results are relegated to the appendix.

4.2 Optimal profit shifting

In the profit shifting scenario, the MNE's optimal choice of λ is given by

$$\lambda = 1 - \exp\left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right). \tag{11}$$

The following lemma provides a formal characterization of how this solution depends on the profit shifting technology, which is governed by the markdown φ, and the potential gain from shifting profits, which is governed by the tax haven's tax rate τ_{i^*}.

Lemma 1 *Under Assumption 1, the share λ of intangible capital sold to the tax haven is:*

1. *decreasing in φ with elasticity given by*

$$\varepsilon_{\varphi}^{\lambda} = -\left(\frac{1-\lambda}{\lambda}\right)\left(\frac{\tau_i - \tau_{i^*}}{1-\tau_i}\right)\varphi < 0, \tag{12}$$

and is equal to zero if $\varphi = 1$;

2. *decreasing in τ_{i^*} with elasticity given by*

$$\varepsilon_{\tau_{i^*}}^{\lambda} = -\left(\frac{1-\lambda}{\lambda}\right)\left(\frac{1-\varphi}{1-\tau_i}\right)\tau_{i^*} < 0. \tag{13}$$

The first part of this lemma says that the smaller the markdown below the competitive price (i.e. the larger is φ), the smaller the fraction of intangible capital that is shifted to the tax haven. In particular, if the MNE has to sell the rights to intangible capital at the competitive price with no markdowm (i.e., $\varphi = 1$), then no profit shifting takes place at all. The second part says that λ is decreasing in the tax haven's tax rate, τ_{i^*}. The elasticity of λ with respect to τ_{i^*} depends on four terms. First, the closer λ is to 1, the larger the reduction. Second, λ is more responsive to τ_{i^*} if the markdown φ is smaller. Third, the elasticity is increasing in the level of the tax rate in the headquarters, τ_i. Finally, it is proportional to τ_{i^*} itself.

4.3 The Effect of Profit Shifting on Intangible Investment

Having characterized the MNE's decision about how much intangible capital to transfer to the tax haven, we can now characterize the effect of this decision on the MNE's intangible investment choice. The optimal intangible capital allocations in the three scenarios are

$$z^{FT} = \left(\frac{\sum_k (1-\tau_k)\,\Lambda_k}{(1-\tau_i)\,p_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}}, \tag{14}$$

$$z^{TP} = \left(\frac{\sum_k \Lambda_k}{p_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}}, \tag{15}$$

$$z^{PS} = z^{TP} \left(1 - \mathcal{C}(\lambda) + \frac{\lambda\,(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}}, \tag{16}$$

where

$$\Lambda_k \equiv \phi \gamma^{\frac{\gamma}{1-\gamma}} p_k^{\frac{1}{1-\gamma}} A_k^{\frac{1}{1-\gamma}} \left(\frac{1}{w_k} \right)^{\frac{\gamma}{1-\gamma}} N_k^{\frac{\phi}{1-\gamma}}. \tag{17}$$

The following proposition summarizes the relationships between these allocations.

Proposition 1 *Under Assumption 1, the following hold:*

1. *if* $\tau_i = \max\{\tau_k\}_{k=1}^K$ *then* $z^{TP} < z^{FT}$*;*
2. $z^{PS} > z^{TP} \iff \varphi < 1$ *and* $z^{PS} = z^{TP} \iff \varphi = 1$*;*
3. z^{PS} *is decreasing in* φ*;*
4. z^{PS} *is decreasing in* τ_{i^*} *with elasticity*

$$\varepsilon_{\tau_{i^*}}^{z^{PS}} = -\left(\frac{1-\gamma}{1-\phi+\gamma} \right) \frac{1}{\left(1 + \frac{1-\mathcal{C}(\lambda)}{\mathcal{C}'(\lambda)} \right)} \left(\frac{\tau_{i^*}}{\tau_i - \tau_{i^*}} \right) < 0. \tag{18}$$

The first part of the proposition states that if the MNE's home country has the highest tax rate across all of the jurisdictions in which the MNE operates, transfer pricing reduces intangible investment, i.e., $z^{TP} < z^{FT}$. Intuitively, requiring foreign affiliates to pay licensing fees to use intangible capital reallocates intangible income to the headquarters, and if the headquarters' income is taxed at a higher rate, the MNE's global profits decline. This demonstrates that asymmetries in tax rates across jurisdictions are more distortionary when MNEs are required to account for intangible income according to the arm's length principle.

The second part of the proposition states that, relative to the transfer pricing scenario, profit shifting increases intangible investment, i.e., $z^{PS} > z^{TP}$, if and only if intangible capital can be sold to the tax haven below the competitive price, i.e., $\varphi < 1$. In this case, as can be seen in (11), $\lambda \in (0,1)$, and we show in the Appendix that this implies the term in parentheses in (16) is strictly greater than one. Intuitively, profit shifting allows the MNE to partially undo the impact of transfer pricing. Transfer pricing forces the MNE to book foreign affiliates' intangible income at the home tax rate, and profit shifting allows the MNE to book some of this income at the tax haven's tax rate instead. In fact, if the the MNE's home country has the highest tax rate, then one can show that $z^{TP} < z^{PS} < z^{FT}$.

The third and fourth parts of the proposition characterize the size of the effect described in the second part. As shown in Lemma 1, the smaller the markdown (the larger φ), the smaller the fraction λ of intangible capital that is sold to the tax haven. This implies that the MNE's profit is decreasing in φ; the closer φ is to the competitive price, the lower the incentive to purchase intangible capital. In turn, this implies that z^{PS} is decreasing in φ. Similarly, z^{PS} is decreasing in the tax haven's tax rate τ_{i^*}. As this rate increases, λ falls, and with it falls the extra gain from intangible investment relative to the transfer pricing scenario. The elasticity of this margin is negative and given by 18. It a product of three terms: (i) technological parameters; (ii) the profit shifting cost function; and (iii) the difference between the tax rates in the tax haven and the MNE's home country. These comparative statics are also illustrated in figure 1.

These results are crucial for understanding the central economic tradeoff we uncover in this paper: profit shifting erodes high-tax countries' tax bases, but also boosts economic activity by increasing MNEs' intangible investment. This tradeoff has important implications for the OECD/G20 BEPS framework. Specifically, a global minimum corporate income tax—which in this simple environment acts like an increase in the tax haven's tax rate—will reduce profit shifting, but this reduction will at the cost of lower economic performance.

4.4 The Effect of the Profit Allocation Rule

We can also use our theory of profit shifting to illustrate the impact of the first pillar of the OECD/G20 framework, which allocates the rights to tax a portion of an MNE's global profits to the regions in which it operates in proportion to these regions' shares of the MNE's overall sales. Under this rule, the tax base of a subsidiary in region k is the sum of local routine profit π_k^r, a share $(1-\theta)$ of local residual profit π_k^R, and a fraction of total global residual profit Π^R that is based on this region's share of the MNE's total global sales:

$$T_k = \pi_k^r + (1-\theta) \cdot \pi_k^R + \theta \cdot \frac{p_k y_k}{\sum_k p_k y_k} \cdot \Pi^R. \tag{19}$$

Routine profit is defined as fraction μ of the revenues in jurisdiction k,

$$\pi_k^r = \mu p_k y_k, \tag{20}$$

and residual profit is defined as the complementary fraction,

$$\pi_k^R = \pi_k^{PS} - \pi_k^r. \tag{21}$$

Global residual profit is the sum of residual profits across regions:

$$\Pi^R = \sum_i \pi_i^R. \tag{22}$$

The two key parameters are: (i) the fraction of residual profits that are allocated across regions based on sales, θ; and (ii) the routine profitability margin, μ. Under the OECD/G20 proposal, these are set accordingly to $\theta = 0.25$ and $\mu = 0.1$, but in what follows we will analyze comparative statics with respect to their values.

Consider now the MNE's modified profit-maximization problem in the profit-shifting scenario under the profit allocation rule:

$$\max_{z^{PS}, \left\{\ell_k^{PS}\right\}_{k=1}^{I}, \lambda} \sum_{k=1}^{I} \left(\pi_k^{PS} - \tau_k T_k\right). \tag{23}$$

The share of intangible capital that is sold to the tax haven is now given by

$$\hat{\lambda} = 1 - \exp\left(-\frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1 - ((1-\theta)\tau_i + \theta\widehat{\tau})}\right). \tag{24}$$

where $\widehat{\tau}$ is the sales-weighted average tax rate across regions:

$$\widehat{\tau} \equiv \sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k}. \tag{25}$$

The MNE's optimal choice of intangible capital is given by

$$\hat{z}^{PS} = \hat{z}^{TP}\left(1 - \mathcal{C}(\lambda) + \frac{(1-\theta)\lambda(1-\varphi)(\tau_i - \tau_{i^*})}{(1 - ((1-\theta)\tau_i + \theta\widehat{\tau}))}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}. \tag{26}$$

We can now establish the following lemma characterizing how $\hat{\lambda}$ depends on the parameters of the profit allocation rule and how it differs from the share λ that is transferred under the existing tax regime.

Lemma 2 *Under Assumption 1, the following hold:*

1. *the fraction of intangible capital sold to the tax haven under the profit allocation rule*

is smaller than under the current regime, i.e., $\hat{\lambda} < \lambda$*;*

2. $\hat{\lambda}$ *is decreasing in* θ *with elasticity given by*

$$\varepsilon_{\theta}^{\hat{\lambda}} = -\mathcal{C}'(\lambda)\left(\frac{1-\lambda}{\lambda}\right)\left(\frac{\theta}{1-\theta}\right)\frac{(1-\widehat{\tau})}{1-((1-\theta)\tau_i+\theta\widehat{\tau})} < 0; \tag{27}$$

3. $\hat{\lambda}$ *is decreasing in* τ_i*, and if the MNE's sales in the tax haven the tax haven are sufficiently small, then*

$$\left|\varepsilon_{\tau_{i^*}}^{\hat{\lambda}}\right| < \left|\varepsilon_{\tau_{i^*}}^{\lambda}\right|. \tag{28}$$

The first part of the lemma establishes that less intangible capital is transferred to the tax haven under the profit allocation rule than under the existing tax regime. This can be seen by comparing 11 and 24. The second part shows that $\hat{\lambda}$, is a decreasing function of the fraction of residual profits allocated based on sales, θ. Finally, the third part shows that $\hat{\lambda}$ is a decreasing function of the tax haven's tax rate, τ_{i^*}, just like λ. Importantly, however, if the tax haven accounts for a sufficiently small share of the MNE's global sales—which is the relevant case in our quantitative analysis—$\hat{\lambda}$ is less responsive to τ_{i^*} than λ. This implies that the profit allocation rule dampens the effect of the second OECD/G20 pillar, the global minimum corporate income tax.

We are now ready to characterize how the profit allocation rule affects the MNE's intangible investment decision.

Proposition 2 *Under Assumption 1, the following hold:*

1. *the allocation of intangible capital under the profit allocation rule, for any* $0 < \theta \leq 1$*, is smaller than under the current regime, i.e.* $\widehat{z}^{PS} < z^{PS}$*;*

2. $\hat{z}^{PS}$ *is decreasing in* θ *with the elasticity given by:*

$$\varepsilon_{\theta}^{\hat{z}^{PS}} = \varepsilon_{\theta}^{\hat{\lambda}}\left(\frac{\hat{\lambda}^2}{1-\hat{\lambda}}\right)\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\left(\widehat{z}^{PS}\right)^{\frac{1-\phi-\gamma}{1-\gamma}} < 0; \tag{29}$$

3. $\hat{z}^{PS}$ *is decreasing in* τ_{i^*}*, and if the MNE's sales in the tax haven are sufficiently small then*

$$\left|\varepsilon_{\tau_{i^*}}^{\hat{z}^{PS}}\right| < \left|\varepsilon_{\tau_{i^*}}^{z^{PS}}\right|. \tag{30}$$

The first part of the proposition states that the profit allocation rule will reduce intangible investment relative to the current regime, i.e., $\widehat{z}^{PS} < z^{PS}$. It can be seen by comparing the solution for z^{PS} in (15) with the solution for $\hat{z}^{PS}$ in (26). The second part states that intangible investment is decreasing in the fraction of residual profits allocated based on sales, θ. The elasticity of this margin is given by (29). It is proportional to the elasticity of $\hat{\lambda}$ with respect to θ given by (27), which itself is negative as shown in the previous lemma. Finally, the third part of the proposition states that intangible investment under the profit allocation rule is decreasing in the tax haven's tax rate, which is also true under the current regime. However, as with the share of intangible capital sold to the tax haven, the size of this effect is smaller under the profit allocation rule, provided that the tax haven is sufficiently small. These comparative statics are illustrated in figure 1.

These findings reveal an important interaction between two OECD/G20 pillars and provide a deeper understanding of the trade-offs that policymakers face. On the one hand, the profit allocation rule decreases profit shifting. On the other hand, although it decreases intangible investment, it also alleviates the negative impact of the global minimum tax. As we will see, these margins play important roles in our quantitative analysis, which we take up in the next sections of the paper.

5 Quantitative Model

In order to assess the macroeconomic implications of our theory of profit shifting, we integrate it into a general equilibrium model with heterogeneous firms in the tradition of the international economics literature. Our quantitative framework synthesizes Helpman, Melitz, and Yeaple (2004) and McGrattan and Prescott (2010). There are I "productive" regions, each populated by a representative household, a measure of heterogeneous firms, and a government. Regions, indexed by i and j, differ in population, total factor productivity, trade costs, FDI costs, and corporate income taxes. Firms in each region decide: where to export and where to open foreign subsidiaries; how much labor to hire in the parent division and each subsidiary; and how much intangible capital to produce in the parent division. Intangible capital is nonrival and is used simultaneously in all of a firms' divisions.

As in section 4, multinational firms (firms with foreign affiliates) use transfer pricing to allocate the costs of producing intangible capital across their foreign affiliates in proportion to the scale at which these affiliates use this capital. Affiliates license the right to use intangible capital from the division that owns this capital, and MNEs can shift profits by selling their intangible capital to affiliates in lower-tax regions. We denote the "productive" region with the lowest corporate income tax rate by denoted by LT. Additionally, there is an "unproductive" tax haven that is populated by a representative household and a government, labelled as TH, where no economic activity takes place. MNEs based in high-tax regions can transfer their intangible capital rights to either (or both) of these regions, provided that they have established affiliates there.

5.1 Households

Each region i has a representative household with preferences over consumption, C_i, and labor supply, L_i, given by

$$u\left(\frac{C_i}{N_i}, \frac{L_i}{N_i}\right) = \log\left(\frac{C_i}{N_i}\right) + \psi_i \log\left(1 - \frac{L_i}{N_i}\right). \tag{31}$$

Households choose consumption and labor supply to maximize utility subject to a budget constraint

$$P_i C_i = W_i L_i + D_i + T_i, \tag{32}$$

where W_i is the wage, D_i is the aggregate dividend payment from firms based in region i, and T_i is a transfer from the government.

Consumption is a constant-elasticity-of-substitution aggregate of products from different source countries,

$$C_i = \left[\sum_{j=1}^{J} \int_{\Omega_{ji}} q_{ji}(\omega)^{\frac{\rho-1}{\rho}} d\omega\right]^{\frac{\rho}{\rho-1}}, \tag{33}$$

where $q_{ji}(\omega)$ is the quantity of variety ω from region j, Ω_{ji} is the set of goods from j available in i (which is determined by firms' exporting and FDI decisions that we will specify in detail later), and ρ is the elasticity of substitution between varieties. The demand curve for each variety can be written as

$$p_{ji}(\omega) = P_i C_i^{\frac{1}{\rho}} q_{ji}(\omega)^{-\frac{1}{\rho}}. \tag{34}$$

The aggregate price index is

$$P_i = \left[\sum_{j=1}^{J} \int_{\Omega_{ji}} p_{ji}(\omega)^{1-\rho} d\omega\right]^{\frac{1}{1-\rho}}. \tag{35}$$

5.2 Firms

Each productive region i has a unit measure Ω_i of firms that compete monopolistically as in Melitz (2003) and Chaney (2008). Each firm is associated with a product variety ω. Firms are heterogeneous in productivity, a, which is drawn from a distribution $F_i(a)$. Firms produce their products using labor and intangible capital. Intangible capital, which we denote by z, is nonrival: it is produced in the home country but can be used to produce abroad as well, provided that a firm pays the cost of setting up a foreign affiliate in another productive region. Foreign affiliates pay licensing fees to use intangible capital according to the rules of transfer pricing. Firms can shift the profits associated with these fees to the low-tax region and/or the tax haven by transferring the rights to intangible capital to affiliates in these regions. Profit shifting is costly, however, and the more capital that is transferred, the larger the cost. Throughout this subsection, we index firms by their productivities instead of their varieties to economize on notation; all firms from a given region with the same productivity make the same decisions.

Production. A firm from region i with productivity a and intangible capital z can produce its good in any productive region j using the technology

$$y_{ij} = \sigma_{ij} A_j a \left(N_j z\right)^{\phi} \ell_j^{\gamma}. \tag{36}$$

This technology is the same as in the theory developed in section 4 with two modifications: it depends on the firm's idiosyncratic productivity as well as region j's aggregate productivity; and the firm's ability to deploy its productivity and intangible capital abroad may be limited by FDI barriers, σ_{ij}, as in McGrattan and Waddle (2020). We assume that $\sigma_{ij} \in [0, 1]$ and that $\sigma_{ii} = 1$.

Research & development. Firms hire workers in their domestic parent corporations to produce intangible capital. We assume that labor productivity in R&D is the same as TFP in production. In other words, it takes $1/A_i$ workers in region i to produce one unit of intangible capital, i.e., the cost to produce z units of intangible capital is $W_i z/A_i$. Following McGrattan and Waddle (2020), we assume that R&D expenditures are tax-deductable.[13]

Trade and foreign direct investment. Firms can sell in the domestic market freely, but serving foreign markets is costly. There are two options for serving foreign markets: pay a fixed cost κ_i^X to export domestically produced goods; or pay a fixed cost κ_i^F to open a foreign affiliate and produce locally. Fixed costs are denominated in units of the home country's labor. Each unit of goods shipped abroad incurs an iceberg transportation cost ξ_{ij}. Firms can simultaneously export to, and produce locally for, the same foreign country; exports and locally-produced products are considered distinct varieties as in Garetto et al. (2019) and

[13] Alternatively, one could treat R&D like tangible investment, but this would require a dynamic model in which a depreciation allowance is deducted from taxes instead. We leave a dynamic treatment of profit shifting for future research.

McGrattan and Waddle (2020).[14] Let $J_X \subseteq I \setminus \{i\}$ denote the set of foreign regions to which a firm exports, and let and $J_F \subseteq I \setminus \{i\}$ denote the set of regions in which it operates a foreign affiliate. The firm's resource constraints can then be written as follows:

$$y_{ii} = q_{ii} + \sum_{j \in J_X} \xi_{ij} q_{ij}^X \tag{37}$$

$$y_{ij} = q_{ij}, \ j \in J_F \tag{38}$$

where we distinguish exported goods, denoted as q_{ij}^X, from goods that are produced and consumed in the same location, q_{ij}.

Transfer pricing. As in section 4, foreign subsidiaries pay licensing fees to use intangible capital. As before, the licensing fee of a subsidiary in region j is given by $\vartheta_{ij} z$, where $\vartheta_{ij} \equiv \gamma y_{ij}/z$ is the marginal revenue product of intangible capital, and the total amount of licensing fees across the conglomerate is $\nu_i z \equiv \sum_{j \in J_F \cup \{i\}} \vartheta_{ij} z$. Note again that this includes the licensing fee for the parent corporation's use of its own intangible capital.

Profit shifting. Also as in section 4, a firm based in a high-tax region can shift its profits by transferring ownership of its intangible capital its affiliates in lower-tax jurisdictions (provided that the firm has paid the fixed costs to establish these affiliates). Consider a firm that sells a fraction λ^{LT} of its intangible capital to the low-tax region and a fraction λ^{TH} to the tax haven. Its affiliate in the former collects licensing fees of

$$\lambda_{LT} \sum_{j \in J_F \cup \{i\} \setminus \{LT\}} \vartheta_{ij} z, \tag{39}$$

while its affiliate in the latter collects

$$\lambda_{TH} \sum_{j \in J_F \cup \{i\}} \vartheta_{ij} z. \tag{40}$$

The domestic parent collects the remaining fees:

$$(1 - \lambda_{LT} - \lambda_{TH}) \sum_{j \in J_F} \vartheta_{ij} z. \tag{41}$$

Setting up an affiliate in the tax haven requires a fixed cost κ_i^{TH}. The cost (paid in units of labor in the home country) to sell a fraction λ of intangible capital to another country

[14] We have also studied a version of the model in which firms must choose whether to export or produce locally for each foreign market (the standard proximity-concentration tradeoff), and the results are similar.

is specified as $\mathcal{C}_{ij}(\lambda) = [\lambda + (1-\lambda)\log(1-\lambda)]\,\psi_{ij}$, where ψ_{ij} governs the marginal cost. Note that in the simple theoretical model in section 4, the profit shifting technology is governed by the discount φ at which the MNE can sell its intangible capital to its affiliate in the tax haven. In our quantitative setting, the marginal cost parameter ψ_{ij} captures this discount as well as the resource cost of shifting profits. This allows for more flexibility and allows our model to generate a wider range of profit-shifting outcomes.[15] The cost to sell a fraction λ_{LT} of intangible capital to the low-tax region and a fraction λ_{TH} to the tax haven is $\mathcal{C}_{i,LT}(\lambda_{LT}) + \mathcal{C}_{i,TH}(\lambda_{TH})$.

5.3 Firm's problem

The firm's objective is to maximize its dividend payout. We describe the firm's problem in three steps: first, in a standard environment without transfer pricing or profit shifting; second, with transfer pricing but without profit shifting; and third, with profit shifting.

5.3.1 Free transfer scenario

Here, the firm chooses where to export (J_X); where to open a foreign affiliate (J_F); how much intangible capital to produce (z); how much to labor to hire in each of its divisions (ℓ_{ij}); and how much to sell to each of its markets (q_{ij}, q_{ij}^X). We can break this problem into two stages, working backwards. In the second stage, the firm maximizes each division's gross operating profits taking J_X, J_F, and z as given. The domestic parent corporation's profits are:

$$\pi_i^D(a,z;J_X) = \max_{q_{ii},\{q_{ij}^X\}_{j\in J_X},\ell_i} \left\{ p_{ii}(q_{ii})q_{ii} + \sum_{j\in J_X} p_{ij}(q_{ij}^X)q_{ij}^X - W_i\ell_i \right\}$$
$$\text{s.t} \quad q_{ii} + \sum_{j\in J_X} \xi_{ij}q_{ij} = y_i = A_i a (N_i z)^{\gamma} \ell_i^{\phi}. \tag{42}$$

Foreign subsidiaries' profits are

$$\pi_{ij}^F(a,z) = \max_{q_{ij},\ell_j} \; p_{ij}(q_{ij})q_{ij} - W_j\ell_j, \; j \in J_F. \tag{43}$$

Note that these objects will not change when we incorporate transfer pricing and profit shifting.

[15] ψ_{ij} enters the first-order conditions for λ and z in exactly the same way as φ, so we could not separately identify the two parameters if they were both included. For example, the solution for λ_j is now given by $\lambda_j = 1 - \exp\left(-\frac{1-\varphi}{\psi_{ij}}\frac{\tau_i - \tau_{i*}}{1-\tau_i}\right)$. We could pick any value for φ and recalibrate ψ_{ij} to match our target moments, and the equilibrium would always be identical.

In the first stage, the firm chooses J_X, J_F, and z to maximize its global net profits, taking into account the cost of producing intangible capital, as well as the fixed costs of exporting and opening foreign affiliates:

$$d_i^{FT}(a) = \max_{z,J_X,J_F} \Bigg\{ (1-\tau_i) \Bigg[\overbrace{\pi_i^D(a,z;J_X) - W_i \Bigg(z/A_i + \sum_{J \in J_X} \kappa_{ijX} + \sum_{j \in J_F} \kappa_{ijF} \Bigg)}^{\pi_{ii}^{FT}} \Bigg] + \sum_{j \in J_F} (1-\tau_j) \underbrace{\pi_{ij}^F(a,z)}_{\pi_{ij}^{FT}} \Bigg\} \tag{44}$$

See Appendix B.1 for more details on the solution to this problem. We use π_{ii}^{FT} and π_{ij}^{FT} to denote the firm's taxable profits in its domestic parent division and foreign subsidiaries, respectively, in this scenario.

5.3.2 Transfer pricing scenario

Here, the firm makes the same choices as in the free transfer scenario, but it takes into account the licensing fees that its foreign affiliates pay to the parent corporation. The first stage of the firm's problem in this scenario is

$$d_i^{TP}(a) = \max_{z,J_X,J_F} \Bigg\{ (1-\tau_i) \Bigg[\overbrace{\pi_i^D(a,z;J_X) - W_i \Bigg(z/A_i + \sum_{J \in J_X} \kappa_{ijX} + \sum_{j \in J_F} \kappa_{ijF} \Bigg) + \sum_{j \in J_F} \vartheta_{ij}(z) z}^{\pi_{ii}^{TP}} \Bigg] + \sum_{j \in J_F} (1-\tau_j) \Bigg[\underbrace{\pi_{ij}^F(a,z) - \vartheta_{ij}(z) z}_{\pi_{ij}^{TP}} \Bigg] \Bigg\} \tag{45}$$

We make explicit the dependence of the licensing fees on the firm's choice of intangible capital by writing $\vartheta_{ij}(z)$ as a function of z. Different from our simple static framework, firms in our quantitative model internalize the effects of their choices of z on transfer prices. See Appendix B.2 for more details on the solution to this version of the problem. π_{ii}^{TP} and π_{ij}^{TP} denote the firm's taxable profits in its domestic and foreign divisions, respectively, in this scenario. The difference between these objects and their counterparts in the free transfer scenario is intangible capital licensing fees, which increase taxable profits in the parent and reduce them in foreign subsidiaries. As we will see, they will be crucial in defining the amount of lost profits in our model with profit shifting.

5.3.3 Profit shifting scenario

Profit shifting adds an additional decision: how much intangible capital to shift to affiliates in the low-tax region and/or tax haven. This problem can be written as

$$
\begin{aligned}
d_i^{PS}(a) = \max_{z, J_X, J_F, \lambda_{TL}, \lambda_{TH}} \Bigg\{ (1-\tau_i) \Bigg[& \pi_i^D(a, z; J_X) - W_i\Bigg(z/A_i + \sum_{J \in J_X} \kappa_{ijX} + \sum_{j \in J_F} \kappa_{ijF} \Bigg) \\
& + \overbrace{\sum_{j \in J_F} (1-\lambda_{LT}-\lambda_{TH}) \vartheta_{ij}(z) z}^{\text{Licensing fee receipts}} + \overbrace{(\varphi_{iLT}\lambda_{LT} + \varphi_{iTH}\lambda_{TH})\, v_i(z) z}^{\text{Proceeds from selling } z} \\
& - \overbrace{(\lambda_{LT}+\lambda_{TH})\vartheta_{ii}(z) z}^{\text{Licensing fee payments}} - \overbrace{W_i \kappa_{iTH} \mathbb{1}(\lambda_{TH} > 0)}^{\text{Tax haven affiliate cost}} \\
& - \overbrace{W_i(\mathcal{C}_{i,TH}(\lambda_{TH}) + \mathcal{C}_{i,LT}(\lambda_{LT}))\nu_i(z) z}^{\text{Cost of transferring } z} \Bigg] \\
+(1-\tau_{LT}) \Bigg[& \pi_{i,LT}^F(a, z) + \overbrace{\sum_{j \in J_F \cup \{i\} \setminus \{LT\}} \lambda_{LT} \vartheta_{ij}(z) z}^{\text{Licensing fee receipts}} \\
& - \underbrace{\varphi_{iLT}\lambda_{LT} v_i(z) z}_{\text{Cost of buying } z} - \underbrace{(1-\lambda_{LT})\vartheta_{iLT}(z) z}_{\text{Licensing fee payment}} \Bigg] \mathbb{1}_{\{LT \in J_F\}} \\
+(1-\tau_{TH}) \Bigg[& \underbrace{\sum_{j \in J_F \cup \{i\}} \lambda_{TH} \vartheta_{ij}(z) z}_{\text{Licensing fee receipts}} - \underbrace{\varphi_{iTH}\lambda_{TH} v_i(z) z}_{\text{Cost of buying } z} \Bigg] \mathbb{1}_{\{\lambda_{TH} > 0\}} \\
+ \sum_{j \in J_F \setminus \{LT\}} (1-\tau_j) \Bigg[& \pi_{ij}^F(a, z) - \underbrace{\vartheta_{ij}(z) z}_{\text{Licensing fee}} \Bigg] \Bigg\}
\end{aligned} \tag{46}
$$

subject to $\lambda_{LT} + \lambda_{TH} \leq 1$ and $\lambda_{LT} \leq \mathbb{1}_{\{LT \in J_F\}}$. The last inequality simply says that you cannot shift profits to the low-tax region if you do not have an affiliate there. Note that firms in the low-tax region do not choose λ_{LT}, only λ_{TH}. See Appendix B.3 for more details on how to solve this problem. The first term in brackets represents the profits of the parent division, π_{ii}^{PS} in this scenario. The second bracketed term represents the profits of the low-tax affiliate, $\pi_{i,LT}^{PS}$, the third represents the profits of the tax-haven affiliate, $\pi_{i,TH}^{PS}$, and the fourth represents the profits of affiliates in other high-tax regions, π_{ij}^{PS}.[16]

[16] We abstract in our model from the Global Intangible Low Tax Income (GILTI), adopted by the U.S. government in 2017, for two reasons. First, once we take the model to the data (see next section) we treat North America as a single region. Second, according to a scarce literature on GILTI, see Clausing (2020b) and Garcia-Bernardo et al. (2022), it had limited impact on profit shifting conducted by the U.S. multinationals.

5.4 Aggregation and accounting measures

Several national and international accounting measures are required to close the model and compare it to the data. Here, we revert to expressing firms' choices as functions of their varieties (ω) for notational brevity.

Gross domestic product. Nominal GDP is the total value of goods produced in a given region:

$$GDP_i = \sum_{j=1}^{I} \int_{\omega \in \Omega_j, i \in J_F(\omega)} p_{ji}(\omega) y_{ji}(\omega) \, d\omega. \tag{47}$$

We compute real GDP by deflating by the consumer price index P_i defined in (35).

Goods trade. Aggregate goods trade flows are given by

$$EX_i^G = \sum_{j \neq i} \int_{\Omega_i} p_{ij}^X(\omega) \left(1 + \xi_{ij}\right) q_{ij}^X(\omega) \, d\omega, \tag{48}$$

$$IM_i^G = \sum_{j \neq i} \int_{\Omega_j} p_{ji}^X(\omega) \left(1 + \xi_{ji}\right) q_{ji}^X(\omega) \, d\omega. \tag{49}$$

Services trade. As in Guvenen et al. (2022), intangible capital licensing fees enter the national accounts as exports or imports of intellectual property services. High-tax regions' services trade flows are given by

$$EX_i^S = \sum_{j \neq i} \int_{\Omega_i} \left[1 - \lambda_{LT}(\omega) - \lambda_{TH}(\omega)\right] \vartheta_{ij}(\omega) z(\omega) \, d\omega, \tag{50}$$

$$IM_i^S = \sum_{j \neq i} \int_{\Omega_i} \left[\lambda_{LT}(\omega) + \lambda_{TH}(\omega)\right] \vartheta_{ij}(\omega) z(\omega) \, d\omega + \sum_{j \neq i} \int_{\Omega_j} \vartheta_{ji}(\omega) z(\omega) \, d\omega. \tag{51}$$

The low-tax region's services trade flows are

$$EX_{LT}^S = \sum_{j \neq i} \int_{\Omega_i} \left[1 - \lambda_{TH}(\omega)\right] \vartheta_{ij}(\omega) z(\omega) \, d\omega + \sum_{j \neq i} \int_{\Omega_j} \lambda_{LT} \vartheta_{ji}(\omega) z(\omega) \, d\omega, \tag{52}$$

$$IM_{LT}^S = \sum_{j \neq i} \int_{\Omega_i} \lambda_{TH}(\omega) \vartheta_{ij}(\omega) z(\omega) \, d\omega + \sum_{j \neq i} \int_{\Omega_j} [1 - \lambda_{LT}(\omega)] \vartheta_{ji}(\omega) z(\omega) \, d\omega. \tag{53}$$

Note that in the transfer-pricing scenario, $\lambda_{LT}(\omega) = \lambda_{TH}(\omega) = 0$. We can also write the tax haven's services exports (it has no imports because foreign affiliates located there do not produce anything) as

$$EX_{TH}^{S} = \sum_{j=1}^{I} \int_{\Omega_j} \lambda_{TH} \vartheta_{ji}(\omega) z(\omega) \; d\omega. \tag{54}$$

Net factor receipts and payments. Net factor receipts from (payments to) foreigners are the sum total of the dividends paid by foreign subsidiaries of domestic multinationals (domestic subsidiaries of foreign multinationals):

$$NFR_i = \sum_{j \neq i} \int_{\Omega_i} (1 - \tau_j) \pi_{ij}^{PS}(\omega) \; d\omega, \tag{55}$$

$$NFP_i = \sum_{j \neq i} \int_{\Omega_j} (1 - \tau_i) \pi_{ji}^{PS}(\omega) \; d\omega. \tag{56}$$

In the free-transfer and transfer-pricing scenarios, we use π_{ij}^{FT} and π_{ij}^{TP}, respectively, to calculate these objects.

Shifted profits. We define the profits shifted out of region j by a firm ω that is based in region i by comparing the profits the firm books in j in the profit-shifting scenario to the profits it would book in the transfer-pricing scenario:

$$\tilde{\pi}_{ij}(\omega) = \pi_{ij}^{TP}(\omega) - \pi_{ij}^{PS}(\omega). \tag{57}$$

When $\tilde{\pi}_{ij}(\omega) > 0$, this indicates that the firm would book more profits in region j in the absence of profit shifting, i.e., the firm has shifted profits away from region j. Aggregating shifted profits by firms at the region level yields the total profits shifted out of region j:

$$\tilde{\Pi}_j = \sum_{i=1}^{I} \int_{\Omega_i} \tilde{\pi}_{ij}(\omega) \; d\omega. \tag{58}$$

Note that $\pi_{ij}^{TP}(\omega)$ is a counterfactual object that can be computed in partial equilibrium or general equilibrium. In partial equilibrium, we calculate it while holding fixed firms' decision rules from the *PS* scenario. In general equilibrium, on the other hand, we re-solve the firm's problem for the *TP* scenario, which changes allocations at the micro level and ultimately at the macro level as well. We use the partial equilibrium version of this measure in our calibration procedure, but we use the general-equilibrium version when analyzing the implications of the two pillars of the OECD proposal.

5.5 Market clearing and equilibrium

In a general equilibrium of our model, the labor market must clear, the government's budget constraint must be satisfied, and the balance of payments must hold in each productive region.

Labor market. Labor demand comes from four sources: production of intermediate goods; production of intangible capital; fixed costs of exporting and setting up foreign affiliates; and the costs of transferring intangible capital. The labor market clearing condition can be written as

$$L_i = \overbrace{\sum_{j=1}^{I} \int_{\Omega_j} \ell_{ji}(\omega)\ d\omega}^{\text{goods production}} + \overbrace{\int_{\Omega_i} z(\omega)/A_i\ d\omega}^{z\text{ production}} + \overbrace{\int_{\Omega_i} \left(\sum_{j \in J_X(\omega)} \kappa_i^X + \sum_{j \in J_F(\omega)} \kappa_i^F + \mathbb{1}_{\{\lambda_{TH}(\omega)>0\}} \kappa_i^{TH} \right) d\omega}^{\text{fixed costs}}$$
$$+ \underbrace{\int_{\Omega_i} \left(\mathcal{C}_{i,TH}(\lambda_{TH}) + \mathcal{C}_{i,LT}(\lambda_{LT}) \right) \nu(\omega) z(\omega)\ d\omega}_{\text{costs of shifting } z} \tag{59}$$

Note that at the macro level, profit shifting diverts labor from goods production and R&D to wasteful administrative costs, potentially offsetting the positive macroeconomic effects of increased R&D at the micro level.

Government budget constraint. We assume that revenue from corporate income taxation is rebated lump-sum to households.[17] In the benchmark profit-shifting model, lump-sum transfers are given by

$$T_i = \tau_i \sum_{j=1}^{I} \int_{\Omega_j} \pi_{ji}^{PS}(\omega)\ d\omega. \tag{60}$$

In the free-transfer and transfer-pricing scenarios, π_{ji}^{FT} and π_{ji}^{TP} are used instead.

Balance of payments. The balance of payments requires that each region's current account must be zero:

$$EX_i^G + EX_i^S - IM_i^G - IM_i^S + NFR_i - NFP_i = 0. \tag{61}$$

Note that several things happen to the balance of payments when a firm shifts profits away from its home region. First, that region's services trade balance worsens: the firm receives

[17] We have also analyzed a version of the model in which labor income taxes adjust to clear the government's budget constraint, and the results are similar.

fewer licensing fees from its foreign subsidiaries and makes more licensing payments. Second, net factor receipts rise: the firm's profits in the tax haven (or low-tax region) rise, and these increased profits are ultimately rebated back to the home country. These two effects offset one another, but not completely: some of the shifted profits are taxed and therefore remain in the tax haven and/or low-tax region. Thus, the net effect is that the current account worsens.[18] To regain equilibrium, that trade balance must improve and/or net factor income balance must improve, which shows up in our model as a real exchange rate depreciation.

Competitive equilibrium. Given a set of parameters and a scenario (free transfer, transfer pricing, or profit shifting), an equilibrium in our model is a set of aggregate prices and quantities $\{W_i, P_i, C_i, L_i\}$ and a set of firm decision rules $\{J_X(\omega), J_F(\omega), z(\omega), \boldsymbol{\ell}(\omega), \boldsymbol{q}(\omega), \lambda_{LT}(\omega), \lambda_{TH}(\omega)\}$ for each productive region $i \in J$ that satisfy

1. the representative household's utility maximization problem described by (31)–(35);
2. the firm's profit-maximization problem described by (42), (43), and either (44), (45), or (46);
3. the labor-market clearing condition (59);
4. the government's budget constraint (60); and
5. the balance of payments (61).

5.6 Calibration

We calibrate our model's parameters so that its equilibrium, given the current international tax regime, reproduces salient facts about production, international trade, foreign direct investment, and, most importantly, profit shifting. Some of the parameters, like elasticities of substitution, are assigned externally to standard values, while others, like population, can be set directly to exact data analogues. The remaining parameters are jointly calibrated by matching a set of target moments. These parameters influence all of the target moments to some degree, but there is one target that provides most of the identification for each parameter. Thus, in what follows, we describe each calibrated parameter alongside its main target. Table 1 lists each parameter in our model alongside its source or target moment. Table 2 provides more detailed information about region-specific target moments and parameter values. Appendix C provides details on the data sources we use to discipline the model.

[18] The reduction in the services trade balance and increase in net factor income is consistent with the accounting of Guvenen et al. (2022). The net negative effect on the balance of payments is consistent with the findings of Hebous et al. (2021).

Regions. We partition the world into five regions. The countries identified as tax havens by Tørsløv et al. (2022) are split into two regions: a low-tax productive region, LT, that includes Belgium, Ireland, Hong Kong, the Netherlands, Singapore, and Switzerland; and an unproductive tax-haven region, TH, which includes Luxembourg, small European countries and territories like Cyprus, Malta, and the Isle of Man, and a number of Carribean countries.[19] The other three regions are North America, Europe (except for the countries in the low-tax and tax-haven regions), and the rest of the world. Data for each region are obtained aggregating or averaging country-level data.

Assigned parameters. The elasticity of substitution between varieties, ρ, is set to the standard value of 5. Each region's population, N_i, is set by aggregating country-level data from the World Bank's World Development Indicators database. Corporate income tax rates, τ_i, are set by averaging country-level estimates of effective corporate income tax rates from Tørsløv et al. (2022).

Technology capital share (ϕ). We set the technology capital share in the production function (36) to match the share of foreign-owned firms' income that accrues to intangible capital, which is estimated by Cadestin et al. (2021) to be 28%. Note that domestic-owned firms have lower intangible income shares, at around 22%. Although we do not target this moment in our calibration, our model is consistent with this fact. This is because technology capital is nonrival, which means that multinational firms have a greater incentive to invest in it than non-MNEs.

Total factor productivity (A_i). Each region's TFP is set to match its aggregated real GDP based on PPP-adjusted data from the World Development Indicators database.

Productivity distribution ($F_i(a)$). We assume that firms' productivities are drawn from Pareto distributions with region-specific tail parameters η_i. We calibrate these tail parameters to match the share of aggregate employment that is accounted for by firms with fewer than 100 times the average number of employees, which is equal to 58.9% in data published by the U.S. Census Bureau. Although this is the only moment of the firm-size distribution that we target, our model's Lorenz curve is very close to its empirical counterpart.

[19]The complete list of countries in the tax-haven region is: Andorra, Anguilla, Antigua, Aruba, the Bahamas, Bahrain, Barbados, Belize, Bermuda, British Virgin Islands, Cayman Islands, Curacao, Cyprus, Gibraltar, Grenada, Guernsey, the Isle of Man, Jersey, Lebanon, Liechtenstein, Luxembourg, Malta, Marshall Islands, Mauritius, Monaco, the Netherlands Antilles, Panama, Puerto Rico, Samoa, Seychelles, Sint Maartin, St. Kitts & Nevis, St. Vincent & the Grenadines, St. Lucia, the Turks & Caicos, and Vanuatu.

Utility weight on leisure (ψ_i). We choose the weight on leisure in the utility function (31) so that the representative household in each country works for one-third of its time endowment, i.e., $L_i = N_i/3$.

Variable trade cost (ξ_{ij}). We set the iceberg trade barriers to match aggregate bilateral imports of goods (agriculture, resource extraction, and manufacturing) relative to nominal GDP. Import data are from the World Input Output Database. Nominal GDP data are from the World Development Indicators. For both, we sum across the countries within each region.

Fixed export cost (κ_i^X). Each region's fixed cost of exporting is chosen so that 22.7% of firms export as reported by Alessandria et al. (2021).

Variable FDI cost (σ_{ij}). We calibrate the parameters that govern the efficiency with which technology capital can be deployed abroad to match the share of each region's gross value added that is accounted for by foreign multinationals. These data come from the OECD AMNE database. This share is equal to 11.12% in North America, 19.82% in Europe, 28.74% in the low-tax region, and 9.55% in the rest of the world.

Fixed FDI cost to productive regions (κ_i^F). The fixed costs of establishing foreign affiliates in other productive regions are set to match the average employment of multinational firms (i.e., firms with foreign affiliates) relative to the overall average employment of all firms. This ratio is equal to 444. The former is calculated using Compustat, while the latter is calculated using data from the U.S. Census.[20]

Variable profit-shifting costs (ψ_{iLT}, ψ_{iTH}). The parameters that govern the cost of transferring technology capital are calibrated by matching Tørsløv et al. (2022)'s estimates of (i) total lost profits, and (ii) the share of lost profits that are shifted to countries in our tax-haven region. As with production and trade data, we obtain region-level measures by summing the country-level estimates reported in this paper. Total lost profits are \$143B for North America, \$216B for Europe, and \$257 billion for the rest of the world. The shares of these totals that are shifted to the tax-haven region are 66.39%, 44.50%, and 71.69%, respectively.

Fixed FDI cost to tax haven (κ_i^{TH}). The fixed costs of establishing affiliates in the tax haven region are set to match the average employment of firms that have affiliates in at least one country in our tax haven region. This ratio is equal to 981. It is also calculated using Compustat.

[20] Compustat contains data on public firms only. We do not have information on employment of private multinational firms. Our approach assumes that private multinationals are similar in size to public multinationals.

5.7 External validation

We have calibrated the key parameters of our model—the profit-shifting costs, ψ_{ij}—to match macroeconomic estimates of aggregate lost profits. However, our calibrated model also matches microeconomic estimates of firm-level profit shifting very closely. We discuss the empirical literature on this topic in detail in Appendix D. The key object of interest in this literature is the semi-elasticity of reported pre-tax profits in an MNE's domestic parent division with respect to the tax differential between the home country and a foreign tax haven. In Table 3 we report three semi-elasticities estimated by key studies in the empirical literature. They range from 0.8 to 1.1, which means that a one percentage point decrease in the tax rate differential—for example, as a consequence of an increase in tax-haven's tax rate—is associated with a 0.8% to 1.1% increase in pre-tax profits reported at home.

To obtain a model counterpart of these elasticities, we estimate the following specification on simulated data generated from counterfactual experiments in which we perturb the different regions' tax rates:

$$\log \pi_i^{k,PS}(\omega) = \beta_0 + \beta_\ell \log \ell_i^k(\omega) + \beta_z \log z^k(\omega) + \beta_\tau \hat{\tau}_i^k + \epsilon_i^k(\omega), \tag{62}$$

where k denotes the index of the counterfactual economy and $\hat{\tau}_i^k$ denotes the tax differential between an MNE's home region and the profit-shifting destination region (either the low-tax region or the tax haven). The parameter of interest is β_τ, which is the relevant semi-elasticity in our model. Appendix D.3 contains more details on how we produce the model-generated data and specify the empirical regression. As Table 3 shows, we obtain an estimate of $\beta_\tau = 0.87$, which lies comfortably within the narrow bounds of the estimates in the empirical literature.

The fact that our calibrated quantitative model is consistent with the microeconomic evidence on profit shifting as well the macroeconomic evidence indicates that it is well-suited to measuring the macroeconomic effects of profit shifting and two-pillar OECD/G20 reform.

6 Quantitative results

Having described the model and its calibration, we turn now to the results of our quantitative analysis. First, we illustrate the effects of transfer pricing and profit shifting by comparing our baseline model to counterfactuals without these ingredients. Second, we analyze the effects of the two pillars of the OECD BEPS project.

6.1 Inspecting the mechanism

Before using the calibrated model to analyze the consequences of changing the global corporate income tax landscape, it is helpful to illustrate the effects of our model's novel ingredients—transfer pricing and profit shifting—under the current tax system. We do this by comparing our baseline model, in which MNEs license technology capital to foreign affiliates according to transfer pricing rules and shift profits by selling technology capital to their affiliates in the tax haven, to two counterfactual models. In the first, the domestic parent corporation retains ownership of technology capital but still licenses this capital to foreign affiliates according to the arm's length principle. We refer to this version as the *no-shifting* counterfactual. In the second, the cost of foreign affiliates' usage of technology capital is not accounted for at all (licensing fees are set to zero). We refer to this version as the *no-transfer-pricing* counterfactual.

Effects of transfer pricing. To illustrate the effects of transfer pricing, panel (a) of table 4 shows how the no-shifting counterfactual compares to the no-transfer-pricing counterfactual. In the highest-tax region in our model, North America, MNEs reduce R&D and produce less output, consistent with part 1 of proposition 1. In other regions, however, MNEs' R&D actually increases. North America, as a large, high-productivity region, is an important FDI destination for these other regions' MNEs. Transfer pricing allows these MNEs to book the returns to intangible capital in their North American subsidiaries, which face the highest tax rates, as profits in their domestic parent divisions, which face lower tax rates. This effect is most pronounced in the low-tax region; this is effectively the reverse of part 1 of proposition 1. In this case, there is also a notable general equilibrium effect for non-MNEs that operates in the opposite direction: greater labor demand by MNEs increases prices, crowding out non-MNEs.

Although the effects of transfer pricing on R&D differ across regions, output falls in equilibrium everywhere, albeit for different reasons. In North America, the decline in output is driven by the response of domestic MNEs. Note that output of foreign MNEs' North American subsidiaries actually rises, but because foreign MNEs account for a relatively small share of overall North American output, this increase is not enough to offset the decline in domestic firms' output. In other regions, the output decline is driven primarily by foreign MNEs, specifically those from North America whose R&D falls.

The effects on corporate tax revenues are heterogeneous across regions. Revenues rise in high-tax North America because licensing fees reallocate income from domestic MNEs' foreign subsidiaries to their parent divisions. In Europe and the rest of the world, revenues fall for the opposite reason: profits of foreign MNEs' subsidiaries in these regions fall when they must pay to use intangible capital. In the low-tax region, revenues rise because domestic MNEs do more R&D and earn more profits globally, which return home in the form of licensing fees.

Effects of profit shifting. Panel (b) of table 4 demonstrates the effects of profit shifting by comparing the baseline model to the no-shifting counterfactual. These effects are easier to explain, as they are the same in the three high-tax regions, North America, Europe, and the rest of the world. In these regions, MNEs increase R&D and produce more output, consistent with part 2 of proposition 1, and this ultimately leads to higher aggregate output. At the same time, profit shifting reduces corporate tax revenues, with the largest effect in Europe.

In the low-tax region, profits shifted in from the high-tax regions amount to almost 4 percent of GDP and tax revenues rise by a full 23.5 percent. In equilibrium, this increase in income raises prices, reducing R&D among both MNEs and non-MNEs. However, the effect of this reduction on aggregate output is offset to a large degree by higher production by foreign MNEs' subsidiaries in this region.

6.2 Policy experiments

We use our calibrated model to conduct four experiments to analyze the macroeconomic consequences of the policies proposed in the OECD/G20 Inclusive Framework on BEPS described above in section 3.2. In the first experiment we focus on the first pillar of this framework, which allocates a portion of an MNE's overall global profit to its subsidiaries based on these subsidiaries' revenues. The second experiment focuses on the second pillar, which imposes a global minimum corporate income tax rate of 15 percent. In the third experiment, we analyze the combined effects of these two pillars together. In the fourth experiment (which is really a set of sub-experiments) we study the combined effects of both pillars under different values for the profit reallocation share and global minimum tax rate. In all four experiments, we restrict attention to long-run analysis, comparing the steady state under the current regime to the steady state after the policy is implemented. Table 5 and figure 2 show the results of these experiments.

OECD Pillar 1: revenue-based profit allocation. As described in section 3.2, the first pillar of the OECD BEPS project allocates, for the purposes of taxation, a fraction of a firm's global profits to the countries in which the firm sells its products. Following the OECD proposal, this allocation is based on these countries' shares of the firm's overall global sales. Importantly, this allocation is independent of whether the firm has a physical presence in these countries, which implies that non-MNE exporters are also subject to this rule. The firm's problem under this rule can be written as

$$d_i^{TP}(a) = \max_{z, J_X, J_F} \Bigg\{ \pi_i^D(a, z; J_X) - W_i \Bigg(z/A_i + \sum_{J \in J_X} \kappa_{ijX} + \sum_{j \in J_F} \kappa_{ijF} \Bigg) + \sum_{j \in J_F} \vartheta_{ij}(z) z - \tau_i T_{ii}(a, z) + \sum_{j \in J_F} \Big[\pi_{ij}^F(a, z) - \vartheta_{ij}(z) z - \tau_j T_{ij}(a, z) \Big], \Bigg\} \tag{63}$$

where $T_{ij}(a, z)$ represents the tax base for region j under the profit allocation rule. In Appendix B.4, we show that $T_{ij}(a, z)$ is given by

$$T_{ii}(a, z) = (1 - \theta) \cdot \pi_i^D(a, z; J_X) + \theta \cdot \frac{R_{ii}(a, z)}{\sum_{k \in \{i\} \cup J_X \cup J_X} R_{ik}(a, z)} \cdot \sum_j \pi_i^D(a, z; J_X), \tag{64}$$

$$T_{ij}(a, z) = (1 - \theta) \cdot \pi_{ij}^F(a, z) + \theta \cdot \frac{R_{ij}(a, z)}{\sum_{k \in \{i\} \cup J_X \cup J_X} R_{ik}(a, z)} \cdot \sum_j \pi_j^F(a, z), \tag{65}$$

where θ is the fraction of residual profits that are reallocated and $R_{ij}(a, z) / \sum_k R_{ik}(a, z)$ represents region j's share of the firm's total global sales.

Panel (a) of table 5 shows the effects of this pillar. It would indeed make a large dent in international profit shifting and materially raise high-tax countries' corporate income tax revenues. Lost profits would fall by 34–40% in North America, Europe and the rest of the world, and tax revenues would increase by 1.6–2.6%. In the low-tax region, profits shifted inward would fall by 31% and tax revenues would fall by 11.4%. At the same time, however, this pillar would decrease output globally. MNEs based in all three high-tax regions would reduce R&D and produce less output, and although non-MNEs would expand slightly in equilibrium, overall output in these regions would decline. The effects would be largest in North America, where MNEs' R&D would fall by 0.8% and aggregate output would fall by 0.13%. In the low-tax region, domestic MNEs would increase R&D, but the decline in foreign MNEs' output in this region would ultimately drag overall output downward as well.

OECD Pillar 2: Global minimum corporate income tax. The second pillar is a global minimum corporate income tax. Following the OECD guidance, we implement this policy through top-up taxes levied by the governments of MNEs' home countries. Specifically, if a firm based in jurisdiction i reports profits in a jurisdiction j where the tax rate is below the global minimum tax rate $\underline{\tau}$, such profits are taxed in jurisdiction i at a rate equal to the tax differential, $\underline{\tau} - \tau_j$. Thus, the additional revenue for jurisdiction i is then

$$\tilde{R}_i = \sum_{j=1}^{I} \int_{\Omega_i} \max\left(\left(\underline{\tau} - \tau_j\right), 0\right) \pi_{ij}^{PS}(\omega)\ d\omega. \tag{66}$$

and then the adjusted budget constraint of the government becomes

$$T_i = \tau_i \sum_{j=1}^{I} \int_{\Omega_j} \pi_{ji}^{PS}(\omega)\ d\omega + \tilde{R}_i. \tag{67}$$

The rest of the equilibrium conditions stay unchanged. Panel (b) of table 5 shows the effects of the second pillar. This policy has even larger effects on high-tax countries' lost profits and tax revenues than the first pillar. Lost profits in North America, Europe, and the rest of the world would fall by 63–85% and tax revenues would rise by 2.6–4.9%. On the other hand, the macroeconomic effects would be smaller. Although European MNEs and MNEs from the rest of the world would reduce R&D by more, North American MNEs' R&D would fall less, and low-tax MNEs' R&D would rise more. The net effect would be negligible effects on GDP in all four regions. Finally, note in the low-tax region, profits shifted inward from the high-tax regions would fall by 51% and corporate income tax revenues would fall by 9.7%, but there would be little effect on aggregate output. This is due to the fact that while domestic firms would actually increase R&D slightly, output produced by foreign MNEs in this region would fall.

Both pillars combined. Panel (c) of table 5 shows the effects of implementing both pillars simultaneously. Consistent with proposition 2, the effects are larger than in either of the first two experiments, but not much larger. Profit shifting would be mostly eliminated: lost profits would fall by 77% in North America, 82% in Europe, and 90% in the rest of the world. Corporate income tax revenues would rise more than under either pillar alone, especially in North America. In the low-tax region, profits shifted inward would fall by 67% and tax revenues by 16.5%. The macroeconomic effects would be slightly larger than under pillar 1 in all regions.

Varying the reallocation share and minimum tax rate. Figure 2 shows how the effects of the two pillars change when their parameters are varied. The x-axis in each plot is pillar 1's profit reallocation share and the y-axis is pillar 2's global minimum tax rate. The first column of plots in the figure shows how the effects on lost profits change and the second column shows how the effects on output change. In both columns, darker shades of red indicate “worse” outcomes: smaller reductions in lost profits in the first column and larger output losses in the second column. The results of this analysis clearly show that a global minimum tax rate is better policy than profit reallocation. Both pillars are effective at reducing profit shifting, but profit reallocation causes much larger output losses. A 17 percent minimum tax rate would essentially eliminate profit shifting entirely but would not reduce output much more than the benchmark 15 percent rate. It would take a profit reallocation share of 90 percent or greater to achieve the same reduction in lost profits, but the output losses from this policy would be an order of magnitude greater.

6.3 Sensitivity Analysis

Our quantitative results are robust to a variety of alternative assumptions and calibrations. Here, we describe the results of three sensitivity analyses that illustrate the impact of some of the most important elements of our model and policy experiments. First, we analyze an alternative profit-reallocation rule for OECD BEPS pillar 1 based on output shares instead of sales shares. Second, we analyze the role of the intangible capital income share. Last, we explore the sensitivity of our results to the costs of shifting profits. Table 6 shows the results of these sensitivity analyses.

Alternative profit reallocation rules. The first pillar of the OECD’s BEPS project reallocates the rights to tax a portion of a firm’s global profits to the regions in which it operates in accordance with these regions’ shares of the firm’s global sales. Importantly, some of these rights are allocated to a firm’s export markets, even if the firm does not operate foreign affiliates in these markets. This aspect of the rule increases effective tax rates for firms based in Europe, the low-tax region, and the rest of the world because North America, which is a large, rich export market, has the highest corporate income tax rate. This reduced these firms’ incentives to invest in intangible capital, even if they do not shift profits at all. This partly explains why pillar 1 has larger macroeconomic consequences than pillar 2, even though the former has smaller effects on profit shifting. To explore the importance of this aspect of pillar 1, we have analyzed the effects of alternative versions in which profit taxation rights are allocated for MNEs only, or are based on output shares instead of sales shares. Panel (a) of table 6 shows the effects of a profit allocation rule than applies only to MNEs, not firms that export but do not operate foreign affiliates. The effects on profit shifting are the same as the OECD’s version but the macroeconomic consequences are smaller, especially outside of North America. Panel (b) of table 6 shows what happens when profit-taxation rights are allocated based on output rather than sales. Under this version of the pillar, export destinations do not receive any taxation rights at all. The results are almost identical to panel (a). These results indicate that allocating taxation rights based on export sales should be avoided.

Intangible share. We have set the share of intangible capital in production, ϕ, to match the share of income that accrues to intangible capital in MNEs' foreign affiliates. This approach ensures that our model captures the extent to which nonrivalry governs MNEs' incentives to invest in intangible capital. This share is the key determinant of the potential scope for profit shifting; a greater intangible share means more licensing fee income that can be transferred to the low-tax region and/or the tax haven. Of course, it is also the key determinant of the macroeconomic impact of policies that affect incentives to invest in intangible capital, including the policies designed to reduce profit shifting that we have studied. Panels (c) and (d) of table 6 show the results of our experiments under alternative calibrations with different intangible shares. In each, we recalibrate all model parameters except for those that govern profit shifting. This allows us to explore how the intangible share affects profit shifting under the current international tax system as well as the effects of changes to this system. The results of these analyses show that a lower intangible share reduces macroeconomic effects of transfer pricing and profit shifting, reduces the amount of profit shifting under the current tax code, and reduces the macroeconomic consequences of the OECD BEPS pillars; the reverse is true for a higher intangible share. However, the extent to which the BEPS pillars reduce profit shifting is about the same as in the baseline model. For example, with a lower intangible share, lost profits in North America fall by 1-0.27/0.45 = 40% under pillar 1, exactly the same as in the baseline model.

Profit-shifting costs. We have set the costs of profit shifting, ψ_{iLT} and ψ_{iTH}, to match estimates in the literature about the amount of profit shifting and the extent to which profits are shifted to low-tax "productive" regions versus "unproductive" tax havens. These estimates are inferred from information about the profitability and labor shares of MNEs' foreign affiliates in these regions—it is impossible to directly measure lost profits without access to detailed information about intra-MNE transactions—so there is some uncertainty about how much profit shifting truly occurs. To determine the sensitivity of our results to these key parameters, we have conducted our experiments in alternative calibrations with in which these parameters are set to higher or lower values. Panel (e) of table 6 shows the results when $\psi_{i,LT}$ and $\psi_{i,TH}$ are halved, while panel (f) shows the results when they are doubled. With lower profit-shifting costs, there is more profit shifting under the current tax system and the OECD BEPS pillars have larger macroeconomic effects; the reverse holds with lower costs.

As in the previous exercise, the BEPS pillars reduce profit shifting by about the same amount as in the baseline. For example, with lower shifting costs, lost profits in North America fall by 1-1.26/2.03 = 38% under pillar 1.

7 Conclusion

We have developed a theory of international profit shifting by multinational enterprises (MNEs) to study the macroeconomic implications of this phenomenon. In our model, MNEs invest in nonrival intangible capital which they can use simultaneously in all of their divisions around the world. MNEs charge their foreign affiliates licensing fees to use intangible capital according to transfer pricing rules, and they can shift profits by transferring the rights to this capital to affiliates in low-tax jurisdictions.

In addition to the methodological contribution that our theory represents, we make two substantive contributions. First, we prove that profit shifting presents a trade-off between economic performance and tax revenues. On the one hand, profit shifting erodes the corporate income tax base in the jurisdiction in which an MNE is based. On the other hand, it incentivizes MNEs to invest in more intangible capital, which boosts output at home as well as abroad. Second, we calibrate our model to match empirical facts about profit shifting under the current international tax regime and use it to quantify the impact of the OECD's plan to eliminate profit shifting. This plan features two pillars: taxing MNEs in the countries in which they sell their products rather than the countries in which they book their profits; and a global minimum corporate income tax rate. We find that this reform would indeed largely eliminate profit shifting and boost tax revenues in high-tax jurisdictions. However, it would also materially reduce intangible capital investment and overall macroeconomic performance.

To put our quantitative results in context, it is helpful to compare them to the effects of other major international policy changes that have been analyzed elsewhere in the literature. Caliendo and Parro (2014) estimate that the North American Free Trade Agreement increased welfare by 0.08% in the United States and reduced it by 0.06% in Canada, while di Giovanni et al. (2014) find that the average country gained 0.13% from liberalizing trade with China. Caliendo et al. (2021) find that the 2004 EU enlargement, which liberalized international labor markets as well as trade, increased welfare in the original EU member states by 0.04%. Despite the small number of firms involved in profit-shifting—far fewer firms engage in multinational production than trade, and only a small fraction of the former shift profits—we find that the macroeconomic effects of the OECD/G20 BEPS framework would be even larger.

We have purposefully left out several aspects of profit shifting and numerous details of the proposed OECD/G20 reform in order to focus on the economic mechanisms at the core of the issue. For example, we have abstracted from manipulation of MNEs' external and internal debt (and the associated interest payments); from optimization of transfer prices; and from the proposed tax rules governing intra-company tax-deducted payments. Also, we have deliberately studied a static economy in which international tax system gives rise only to intratemporal distortions. Thus, we view our results as a lower bound; in a dynamic model, corporate taxes also distort the intertemporal margin. We leave all these important considerations for future research.

References

Alessandria, G., H. Choi, and K. J. Ruhl (2021): "Trade Adjustment Dynamics and the Welfare Gains from Trade," *Journal of International Economics*, 131, Article 103458.

Allingham, M. G. and A. Sandmo (1972): "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, 1, 323–338.

Antrás, P. and S. R. Yeaple (2014): "Multinational Firms and the Structure of International Trade," in *Handbook of International Economics*, ed. by G. Gopinath, . Helpman, and K. Rogoff, Elsevier, vol. 4 of *Handbook of International Economics*, chap. 0, 55–130.

Arkolakis, C., F. Eckert, and R. Shi (2021): "Combinatorial discrete choice," Working Paper.

Auerbach, A. J. (1983): "Corporate Taxation in the United States," *Brookings Papers on Economic Activity*, 14, 451–514.

Barro, R. J. and J. Furman (2018): "Macroeconomic Effects of the 2017 Tax Reform," *Brookings Papers on Economic Activity*, 49, 257–345.

Beer, S., R. de Mooij, and L. Liu (2020): "International Corporate Tax Avoidance: A Review of the Channels, Magnitudes, and Blind Spots," *Journal of Economic Surveys*, 34, 660–688.

Bhandari, A. and E. R. McGrattan (2020): "Sweat Equity in U.S. Private Business*," *The Quarterly Journal of Economics*, 136, 727–781.

Blouin, J. and L. A. Robinson (2020): "Double Counting Accounting: How Much Profit of Multinational Enterprises Is Really in Tax Havens?" SSRN Scholarly Paper ID 3491451, Social Science Research Network, Rochester, NY.

Bolwijn, R., B. Casella, and D. Rigo (2018): "Establishing the baseline: estimating the fiscal contribution of multinational enterprises," *Transnational Corporations*, 25, 111–143.

Cadestin, C., A. Jaax, S. Miroudot, and C. Z urcher (2021): "Multinational Enterprises and Intangible Capital," Tech. Rep. 118.

Caliendo, L., L. D. Opromolla, F. Parro, and A. Sforza (2021): "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement," *Journal of Political Economy*, 129, 3491–3545.

Caliendo, L. and F. Parro (2014): "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82, 1–44.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707–21.

Clausing, K. A. (2016): "THE EFFECT OF PROFIT SHIFTING ON THE CORPORATE TAX BASE IN THE UNITED STATES AND BEYOND," *National Tax Journal*, 69, 905–934.

——— (2020a): "How Big Is Profit Shifting?" SSRN Scholarly Paper ID 3503091, Social Science Research Network, Rochester, NY.

——— (2020b): "Profit Shifting Before and After the Tax Cuts and Jobs Act," *National Tax Journal*.

Corrado, C., C. Hulten, and D. Sichel (2009): "INTANGIBLE CAPITAL AND U.S. ECONOMIC GROWTH," *Review of Income and Wealth*, 55, 661–685.

Crivelli, E., R. A. De Mooij, and M. M. Keen (2015): *Base erosion, profit shifting and developing countries*, International Monetary Fund.

Delis, F., M. Delis, L. Laeven, and S. Ongena (2021): "Global Evidence on Profit Shifting Within Firms and Across Time," CEPR Discussion Papers 16615, C.E.P.R. Discussion Papers.

Devereux, M., A. Auerbach, M. Keen, P. Oosterhuis, W. Schön, and J. Vella (2021): *Taxing profit in a global economy*, Oxford University Press.

Dharmapala, D. (2014): "What Do We Know about Base Erosion and Profit Shifting? A Review of the Empirical Literature," *Fiscal Studies*, 35, 421–448.

di Giovanni, J., A. A. Levchenko, and J. Zhang (2014): "The Global Welfare Impact of China: Trade Integration and Technological Change," *American Economic Journal: Macroeconomics*, 6, 153–83.

Dowd, T., P. Landefeld, and A. Moore (2017): "Profit Shifting of U.S. Multinationals," *Journal of Public Economics*, 148, 1–13.

Dyrda, S., G. Hong, and J. Steinberg (2022): "Intangibles in the World of Tax Competition and Profit Shifting," Mimeo, Department of Economics, University of Toronto.

Ewens, M., R. H. Peters, and S. Wang (2019): "Measuring Intangible Capital with Market Prices," Working Paper 25960, National Bureau of Economic Research.

Garcia-Bernardo, J., P. Janský, and G. Zucman (2022): "Did the Tax Cuts and Jobs Act Reduce Profit Shifting by US Multinational Companies?" Working Paper 30086, National Bureau of Economic Research.

Garetto, S., L. Oldenski, and N. Ramondo (2019): "Multinational expansion in time and space," Tech. rep., National Bureau of Economic Research.

Grubert, H. and J. Mutti (1991): "Taxes, Tariffs and Transfer Pricing in Multinational Corporate Decision Making," *The Review of Economics and Statistics*, 73, 285–293.

Guvenen, F., J. Mataloni, Raymond J., D. G. Rassier, and K. J. Ruhl (2022): "Offshore Profit Shifting and Aggregate Measurement: Balance of Payments, Foreign Investment, Productivity, and the Labor Share," *American Economic Review*, 112, 1848–84.

Harberger, A. C. (1962): "The Incidence of the Corporation Income Tax," *Journal of Political Economy*, 70, 215–240.

Haufler, A. and G. Schjelderup (2000): "Corporate Tax Systems and Cross Country Profit Shifting," *Oxford Economic Papers*, 52, 306–325.

Hebous, S., A. Klemm, and Y. Wu (2021): "How Does Profit Shifting Affect the Balance of Payments?" Tech. Rep. WP/21/14.

Heckemeyer, J. H. and M. Overesch (2017): "Multinationals' profit response to tax differentials: Effect size and shifting channels," *Canadian Journal of Economics/Revue canadienne d'économique*, 50, 965–994.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004): "Export versus FDI with heterogeneous firms," *American Economic Review*, 94, 300–316.

Hines, James R., J. and E. M. Rice (1994): "Fiscal Paradise: Foreign Tax Havens and American Business*," *The Quarterly Journal of Economics*, 109, 149–182.

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

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

Johannesen, N. (2022): "The global minimum tax," *Journal of Public Economics*, 212, 104709.

Johansson, A., Øystein Bieltvedt Skeie, S. Sorbe, and C. Menon (2017): "Tax planning by multinational firms," .

Kaymak, B. and I. Schott (2018): "Corporate Tax Cuts and the Decline of the Labor Share," 2018 Meeting Papers 943, Society for Economic Dynamics.

——— (2019): "Loss-offset provisions in the corporate tax code and misallocation of capital," *Journal of Monetary Economics*, 105, 1–20.

Keen, M. and K. A. Konrad (2013): "Chapter 5 - The Theory of International Tax Competition and Coordination," in *handbook of public economics, vol. 5*, ed. by A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez, Elsevier, vol. 5 of *Handbook of Public Economics*, 257–328.

Koh, D., R. Santaeulàlia-Llopis, and Y. Zheng (2020): "Labor Share Decline and Intellectual Property Products Capital," *Econometrica*, 88, 2609–2628.

McGrattan, E. R. and E. C. Prescott (2010): "Technology Capital and the US Current Account," *American Economic Review*, 100, 1493–1522.

McGrattan, E. R. and A. Waddle (2020): "The Impact of Brexit on Foreign Investment and Production," *American Economic Journal: Macroeconomics*, 12, 76–103.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

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

OECD (2014): *Guidance on Transfer Pricing Aspects of Intangibles.*

——— (2015): *Measuring and Monitoring BEPS, Action 11 - 2015 Final Report.*

——— (2017): *Model Tax Convention on Income and on Capital: Condensed Version 2017.*

——— (2022): *OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations 2022.*

O'Mahony, M., C. Corrado, J. Haskel, M. Iommi, C. Jona-Lasinio, and M. Mas (2018): "Advancements in measuring intangibles for European economies," *EURONA*, 2017.

Peters, R. H. and L. A. Taylor (2017): "Intangible capital and the investment-q relation," *Journal of Financial Economics*, 123, 251–272.

Rotberg, S. and J. B. Steinberg (2022): "Tax Evasion and Capital Taxation," Working paper.

Slemrod, J. and J. D. Wilson (2009): "Tax competition with parasitic tax havens," *Journal of Public Economics*, 93, 1261–1270.

Tiebout, C. M. (1956): "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64, 416–424.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015): "An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production," *Review of International Economics*, 23, 575–605.

Tørsløv, T., L. Wier, and G. Zucman (2022): "The Missing Profits of Nations," *The Review of Economic Studies*.

Wildasin, D. E. (1988): "Nash equilibria in models of fiscal competition," *Journal of Public Economics*, 35, 229–240.

Zodrow, G. R. and P. Mieszkowski (1986): "Pigou, Tiebout, property taxation, and the underprovision of local public goods," *Journal of Urban Economics*, 19, 356–370.

Table 1: Calibration: overview

Parameter	Description	Value(s)	Target/source
(a) Assigned parameters			
ϱ	EoS between products	5	Standard
N_j	Population	Varies	World Development Indicators
τ_j	Corporate income tax rate	Varies	Tørsløv, Wier, and Zucman (2022)
(b) Calibrated parameters			
ϕ	Technology capital share	0.11	MNEs' intangible income share
A_i	Total factor productivity	Varies	Real GDP
η_i	Productivity dispersion	Varies	Large firms' employment share
ψ_i	Utility weight on leisure	Varies	$L_i = N_i/3$
ξ_{ij}	Variable export cost	Varies	Bilateral imports/GDP
κ_i^X	Fixed export cost	Varies	Pct. of firms that export
σ_i	Variable FDI cost	Varies	Foreign MNEs' share of value added
κ_i^F	Fixed FDI cost	Varies	Avg. emp. of firms w/ foreign affiliates
ψ_{iLT}	Cost of shifting profits to LT	Varies	Total lost profits
ψ_{iTH}	Cost of shifting profits to TH	Varies	Share of profits shifted to TH
κ_i^{TH}	Fixed cost of TH affiliate	Varies	Avg. emp. of firms w/ TH affiliates

Table 2: Calibration: details

Region	North America	Europe	Low-tax	RoW	Tax haven
(a) Region-specific target moments					
Population (NA = 100)	100	92	11	1,323	–
Real GDP (NA = 100)	100	80.78	14.57	297.10	–
Corporate tax rate (%)	22.5	17.3	11.4	17.4	3.3
Foreign MNEs' VA share (%)	11.12	19.82	28.73	9.55	–
Total lost profits ($B)	143	216	–	257	–
Lost profits to TH (%)	66.4	44.5	–	71.1	–
Imports from... (% GDP)					
North America	–	1.28	1.77	1.74	–
Europe	1.70	–	12.39	3.78	–
Low tax	0.35	2.98	–	0.59	–
Row	6.15	7.96	6.78	–	–
(b) Internally-calibrated parameter values					
TFP (A_i)	1.00	0.89	1.58	0.20	–
Prod. dispersion (η_i)	4.28	4.31	4.83	4.12	–
Utility weight on leisure (ψ_i)	1.06	1.08	1.09	1.06	–
Fixed export cost (κ_i^X)	1.7e-3	3.5e-3	1.0e-3	1.4e-2	–
Variable FDI cost (σ_i)	0.47	0.56	0.52	0.53	–
Fixed FDI cost (κ_i^F)	1.80	1.59	0.46	8.75	–
Cost of shifting profits to LT (ψ_{iLT})	3.40	0.38	–	2.35	–
Cost of shifting profits to TH (ψ_{iTH})	2.25	1.25	–	1.76	–
Fixed FDI cost to TH (κ_i^{TH})	0.09	0.06	–	0.59	–
Variable trade cost from...					
North America	–	3.21	3.41	2.07	–
Europe	1.89	–	1.69	1.33	–
Low tax	2.04	1.59	–	1.56	–
RoW	2.26	2.59	3.01	–	–

Notes: Population and real GDP from World Bank WDI. Corporate tax rate from Tørsløv et al. (2022). Foreign MNEs' VA share from OECD AMNE database. Fractions of firms with foreign affiliates from Compustat. Lost profits from Tørsløv et al. (2022). Imports/GDP from WIOD. Dashes (–) represent "not applicable."

Table 3: Semi-elasticity of the profit shifting margin: model vs. data

Study	Data source	Headline point estimate
Johansson et al. (2017)	ORBIS, 2000-2010	1.11
Heckemeyer and Overesch (2017)	Meta data: 27 studies, 203 estimates	0.79
Beer et al. (2020)	Meta data: 38 studies, 402 estimates	0.98
This paper	Simulated model data	0.87

Notes: The semi-elasticity of profit shifting represents the effect of a one-percentage-point decrease in the tax rate differential—for example, as a consequence of an increase in the tax haven's tax rate—on the log of pre-tax profits. For Johansson et al. (2017), we report the estimate based on their Table 1. A 1 percentage point tax difference is associated with a 0.069p.p. reduction in the profit-to-assets ratio (Table 1, column 1). The average MNE in the sample has a profit-to-assets ratio of 6.2%. Thus, the effect corresponds to a reduction in profits of $0.069/6.2\% \approx 1.11$ (see their footnote 31). For Heckemeyer and Overesch (2017), we report the consensus estimate provided in their Table 3. For Beer et al. (2020), we report the preferred estimate provided in column 4 of their Table 2. Refer to Appendix D.3 for details of our implementation of the model estimate.

Table 4: Inspecting the mechanism

			Value added (% chg.)				Tech. capital (% chg.)		
Region	Lost profits (% GDP)	Corp. tax rev. (% chg.)	Total	Non MNEs	Domestic MNEs	Foreign MNEs	Total	Non MNEs	Domestic MNEs
(a) Effects of transfer pricing (no transfer pricing vs. no shifting)									
North America	0.00	4.32	-0.16	0.36	-0.85	0.35	-0.54	0.58	-1.34
Europe	0.00	-2.34	-0.17	-0.15	-0.11	-0.31	0.12	0.06	0.17
Low tax	0.00	-2.17	-0.25	-0.72	1.10	-0.56	0.74	-0.75	2.28
Rest of world	0.00	-0.41	-0.18	-0.18	-0.15	-0.31	0.05	0.00	0.08
(b) Effects of profit shifting (no shifting vs. baseline)									
North America	0.68	-3.82	0.08	-0.00	0.15	0.15	0.21	-0.11	0.45
Europe	1.05	-5.43	0.06	-0.02	0.17	0.04	0.26	-0.07	0.55
Low tax	-4.37	23.52	-0.04	-0.33	-0.29	0.64	-0.55	-0.60	-0.49
Rest of world	0.50	-2.59	0.04	-0.01	0.08	0.08	0.12	-0.06	0.27

Table 5: Effects of OECD BEPS pillars

Region	Lost profits (benchmark = 1)	Corp. tax rev. (% chg.)	Value added (% chg.) Total	Value added (% chg.) Non MNEs	Value added (% chg.) Domestic MNEs	Value added (% chg.) Foreign MNEs	Tech. capital (% chg.) Total	Tech. capital (% chg.) Non MNEs	Tech. capital (% chg.) Domestic MNEs
(a) Pillar 1: Profit reallocation									
North America	0.60	2.54	-0.13	-0.01	-0.30	-0.05	-0.40	0.15	-0.80
Europe	0.66	2.61	-0.14	-0.10	-0.18	-0.17	-0.10	0.04	-0.21
Low tax	0.69	-11.40	-0.13	-0.10	0.36	-0.56	0.79	0.23	1.35
Rest of world	0.63	1.63	-0.13	-0.11	-0.15	-0.19	-0.05	0.02	-0.10
(b) Pillar 2: Global minimum tax rate									
North America	0.37	3.24	-0.06	0.01	-0.10	-0.13	-0.15	0.08	-0.31
Europe	0.26	4.89	-0.02	0.04	-0.11	-0.01	-0.22	0.06	-0.45
Low tax	0.49	-9.70	0.02	0.23	0.19	-0.46	0.32	0.36	0.28
Rest of world	0.15	2.64	-0.01	0.04	-0.05	-0.04	-0.11	0.06	-0.24
(c) Pillars 1 & 2 together									
North America	0.23	4.36	-0.17	-0.02	-0.36	-0.11	-0.48	0.17	-0.94
Europe	0.18	5.43	-0.16	-0.09	-0.24	-0.18	-0.21	0.06	-0.43
Low tax	0.33	-16.46	-0.13	0.07	0.50	-0.98	1.00	0.48	1.51
Rest of world	0.10	3.20	-0.14	-0.09	-0.17	-0.21	-0.10	0.05	-0.22

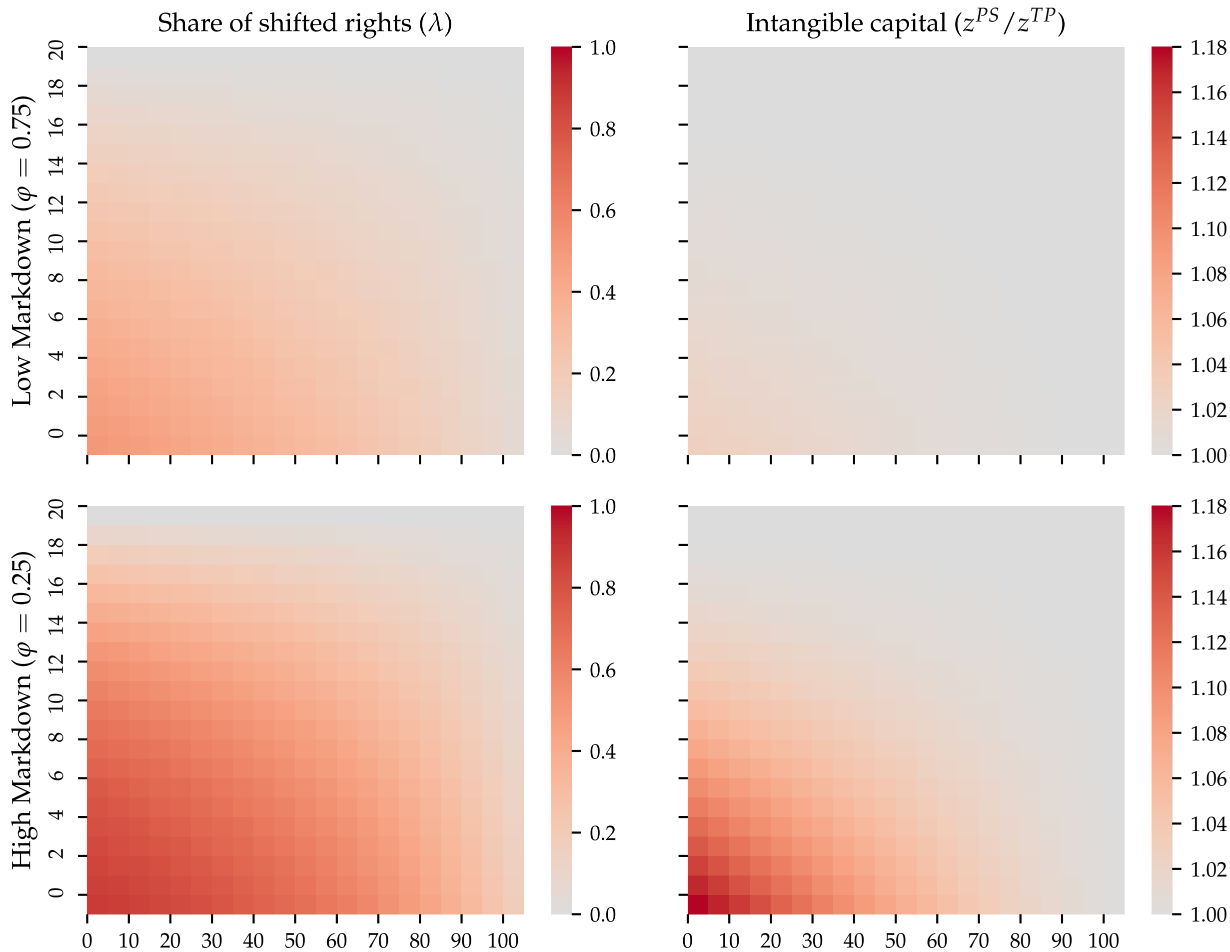
Notes: Lost profits are measured relative to the benchmark. Note that for the low-tax region, lost profits are negative in both the benchmark equilibrium and in the policy counterfactuals, i.e., profits are shifted inward to the low-tax region. However, the magnitude of these lost profits are smaller in the counterfactuals. For example, in panel (b), the amount of profits shifted into the low-tax region under pillar 2 is about half of the amount in the benchmark.

Table 6: Sensitivity analysis

Experiment	Lost profits (benchmark = 1) North America	Europe	Low tax	Rest of World	GDP (% chg.) North America	Europe	Low tax	Rest of World
(a) Profit reallocation rule applies to MNEs only								
Pillar 1	0.60	0.66	0.69	0.63	-0.12	-0.10	-0.06	-0.09
Pillars 1 & 2 together	0.23	0.18	0.33	0.10	-0.16	-0.12	-0.06	-0.10
(b) Production-based profit reallocation rule								
Pillar 1	0.60	0.66	0.69	0.63	-0.12	-0.09	-0.06	-0.09
Pillars 1 & 2 together	0.23	0.18	0.33	0.10	-0.16	-0.11	-0.06	-0.09
(c) Low intangible share								
Effects of transfer pricing	–	–	–	–	-0.06	-0.08	-0.12	-0.08
Effects of profit shifting	0.45	0.45	0.48	0.44	0.02	0.02	-0.04	0.01
Pillar 1	0.27	0.30	0.33	0.28	-0.08	-0.11	-0.12	-0.10
Pillar 2	0.17	0.12	0.23	0.07	-0.02	-0.01	0.02	0.00
Pillars 1 & 2 together	0.10	0.08	0.16	0.04	-0.10	-0.12	-0.13	-0.10
(d) High intangible share								
Effects of transfer pricing	–	–	–	–	-0.27	-0.26	-0.38	-0.28
Effects of profit shifting	1.60	1.60	1.60	1.63	0.16	0.12	-0.04	0.09
Pillar 1	0.96	1.06	1.10	1.03	-0.19	-0.18	-0.14	-0.17
Pillar 2	0.59	0.42	0.78	0.25	-0.11	-0.05	0.02	-0.03
Pillars 1 & 2 together	0.36	0.28	0.53	0.16	-0.25	-0.22	-0.15	-0.18
(e) Low profit-shifting costs								
Effects of profit shifting	2.03	1.96	1.90	2.05	0.15	0.10	-0.09	0.06
Pillar 1	1.26	1.31	1.33	1.32	-0.16	-0.16	-0.10	-0.14
Pillar 2	0.79	0.53	0.96	0.33	-0.10	-0.03	0.06	-0.01
Pillars 1 & 2 together	0.48	0.35	0.66	0.21	-0.23	-0.19	-0.09	-0.15
(f) High profit-shifting costs								
Effects of profit shifting	0.48	0.50	0.51	0.47	0.04	0.04	-0.02	0.03
Pillar 1	0.28	0.32	0.35	0.29	-0.12	-0.13	-0.14	-0.13
Pillar 2	0.17	0.13	0.25	0.07	-0.03	-0.01	0.01	-0.00
Pillars 1 & 2 together	0.10	0.09	0.17	0.04	-0.14	-0.15	-0.15	-0.13

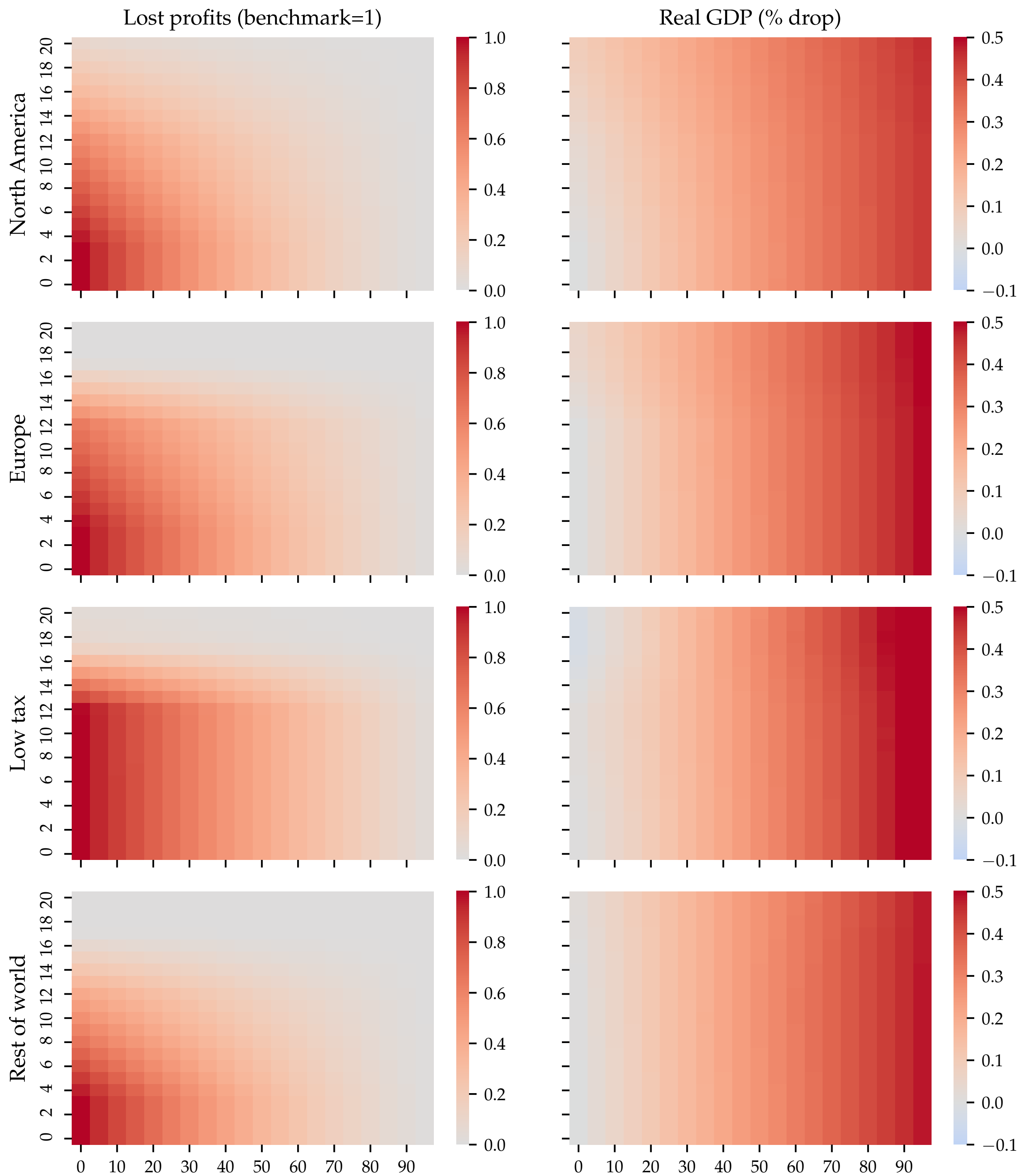
Notes: In panel (a), the profit-reallocation rule for pillar 1 applies only to MNEs (not firms that export but do not operate foreign affiliates). In panel (b), the rule is based on value added rather than sales; profits are not reallocated to export destinations. In panels (c) and (d), the intangible share is changed and all parameters except for profit-shifting costs are recalibrated. Lost profits are measured relative to the benchmark equilibrium in the baseline calibration. In panel (e), the parameters that govern the marginal cost of profit shifting, ψ_{ij}, are halved; in panel (f), they are doubled.

Figure 1: Comparative statics: Shifted property rights and Intangible Capital



Notes: X-axis in each plot represents the reallocation share for pillar 1. Y-axis in each plot represents the global minimum corporate income tax rate for pillar 2. The comparative statics is computed using the corporate income taxes, TFP and populations as in the quantitative model. All prices are normalized to 1. We set $\phi = 0.11$ and $\varphi = 0.64$. The results are presented for North America.

Figure 2: Varying the sizes of the pillars



Notes: Each column reports effects on one variable for each region. First column: Lost profits (reported relative to the benchmark equilibrium). Second column: real GDP (reported as a percent change from the benchmark equilibrium). X-axis in each plot represents the reallocation share for pillar 1. Y-axis in each plot represents the global minimum corporate income tax rate for pillar 2.

Appendix

A Proofs of Analytical Results

This Appendix contains the proofs of the lemmas and propositions from the main body of the paper.

A.1 Main Lemmas

Proof of Lemma 1.

Rewrite the problem 10 using definitions of profits as

$$\max_{z,\lambda,\{l_i\}_{i=1}^{I}} (1-\tau_i)\left(p_i\left(A_i\left(N_i z\right)^{\phi} l_i^{\gamma}\right) - w_i l_i - p_i z + z\left[\varphi\lambda\sum_k \vartheta_k(z) - \lambda\vartheta_i(z) + (1-\lambda)\sum_{k\neq i}\vartheta_k(z) - \sum_k \vartheta_k(z)\mathcal{C}(\lambda)\right]\right)$$
$$+ (1-\tau_{i^*})\left(p_{i^*}\left(A_{i^*}\left(N_{i^*} z\right)^{\phi} l_{i^*}^{\gamma}\right) - w_k l_{i^*} + z\left[\lambda\sum_{k\neq i^*}\vartheta_k(z) - (1-\lambda)\vartheta_{i^*}(z) - \varphi\lambda\sum_k \vartheta_k(z)\right]\right)$$
$$+ (1-\tau_k)\sum_{k\neq i,i^*}\left(p_k\left(A_k\left(N_k z\right)^{\phi} l_k^{\gamma}\right) - w_k l_k - \vartheta_k(z)\, z\right). \tag{68}$$

The FOCs are then:

$$l_i : 0 = \gamma p_i A_i \left(N_i z\right)^{\phi} l_i^{\gamma-1} - w_i, \tag{69}$$

$$z : 0 = \sum_k (1-\tau_k)\phi N_k p_k A_k \left(N_k z\right)^{\phi-1} l_k^{\gamma} + (1-\tau_i)\left[-p_i + \varphi\lambda\sum_k \vartheta_k(z) - \lambda\vartheta_i(z) + (1-\lambda)\sum_{k\neq i}\vartheta_k(z) - \sum_k \vartheta_k(z)\mathcal{C}(\lambda)\right]$$
$$+ (1-\tau_{i^*})\left[\lambda\sum_{k\neq i^*}\vartheta_k(z) - (1-\lambda)\vartheta_{i^*}(z) - \varphi\lambda\sum_k \vartheta_k(z)\right] - \sum_{k\neq i,i^*}(1-\tau_k)\vartheta_k(z), \tag{70}$$

$$\lambda : 0 = (1-\tau_i)\, z\left[\varphi\sum_k \vartheta_k(z) - \vartheta_i(z) - \sum_{k\neq i}\vartheta_k(z) - \sum_k \vartheta_k(z)\mathcal{C}'(\lambda)\right] \tag{71}$$
$$+ (1-\tau_{i^*})\, z\left[\sum_{k\neq i^*}\vartheta_k(z) + \vartheta_{i^*}(z) - \varphi\sum_k \vartheta_k(z)\right].$$

Inspect the FOC wrt to λ:

$$0 = (1-\tau_i)\, z\left[\varphi\sum_k \vartheta_k(z) - \vartheta_i(z) - \sum_{k\neq i}\vartheta_k(z) - \sum_k \vartheta_k(z)\mathcal{C}'(\lambda)\right] + (1-\tau_{i^*})\, z\left[\sum_{k\neq i^*}\vartheta_k(z) + \vartheta_{i^*}(z) - \varphi\sum_k \vartheta_k(z)\right]$$
$$0 = (1-\varphi)\sum_k \vartheta_k(z)\left[\tau_i - \tau_{i^*}\right] - (1-\tau_i)\sum_k \vartheta_k(z)\mathcal{C}'(\lambda),$$

which yields

$$\lambda = (\mathcal{C}')^{-1}\left[(1-\varphi)\,\frac{(\tau_i - \tau_{i^*})}{1-\tau_i}\right]. \tag{72}$$

Under Assumption 1 this can be written as

$$\lambda = 1 - \exp\left(-\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right).$$

Now towards proving the lemma, we have

$$\begin{aligned}
\frac{\partial \lambda}{\partial \varphi} &= -\exp\left(-\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right)\left(\frac{\tau_i - \tau_{i^*}}{1-\tau_i}\right) < 0,\\
\frac{\partial \lambda}{\partial \tau_{i^*}} &= -\exp\left(-\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right)\left(\frac{1-\varphi}{1-\tau_i}\right) < 0,
\end{aligned}$$

and therefore the elasticities are

$$\begin{aligned}
\varepsilon^{\lambda}_{\varphi} = \frac{\partial \lambda}{\partial \varphi}\frac{\varphi}{1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)} &= -\exp\left(-\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right)\left(\frac{\tau_i - \tau_{i^*}}{1-\tau_i}\right)\frac{\varphi}{1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)}\\
&= -\left(\frac{1-\lambda}{\lambda}\right)\left(\frac{\tau_i - \tau_{i^*}}{1-\tau_i}\right)\varphi,
\end{aligned}$$

$$\begin{aligned}
\varepsilon^{\lambda}_{\tau_{i^*}} = \frac{\partial \lambda}{\partial \tau_{i^*}}\frac{\tau_{i^*}}{1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)} &= -\exp\left(-\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right)\left(\frac{1-\varphi}{1-\tau_i}\right)\frac{\tau_{i^*}}{1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)}\\
&= -\left(\frac{1-\lambda}{\lambda}\right)\left(\frac{(1-\varphi)}{1-\tau_i}\right)\tau_i^*,
\end{aligned}$$

which proves 1. and 2. ■

The following lemma will be useful in our further derivations.

Lemma 3 *The allocations of intangible capital are as follows:*

$$\begin{aligned}
z^{FT} &= \left(\frac{\sum_k (1-\tau_k)\,\Lambda_k}{(1-\tau_i)\,p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}, &(73)\\
z^{TP} &= \left(\frac{\sum_k \Lambda_k}{p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}, &(74)\\
z^{PS} &= z^{TP}\left((1-\mathcal{C}(\lambda)) + \frac{\lambda\,(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}. &(75)
\end{aligned}$$

Proof. Free transfer of z requires $\vartheta_k(z) = 0$ thus the 70 becomes

$$z = \left(\frac{\sum_k (1-\tau_k) \Lambda_k}{(1-\tau_i) p_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}},$$

and hence we obtain 73. For the transfer pricing case we have $\lambda = 0$ and the 70 becomes

$$0 = z^{\frac{\phi+\gamma-1}{1-\gamma}} \sum_k (1-\tau_k) \Lambda_k - (1-\tau_i) p_i - \sum_k \vartheta_k(z) (\tau_i - \tau_k),$$

where

$$\begin{aligned}
\vartheta_k(z) &= \phi p_k N_k \left(A_k (N_k z)^{\phi-1} l_k^{\gamma} \right) \\
&= \phi p_k N_k \left(A_k (N_k z)^{\phi-1} \left(\frac{\gamma p_i A_i (N_i z)^{\phi}}{w_i} \right)^{\frac{\gamma}{1-\gamma}} \right) \\
&= \phi \gamma^{\frac{\gamma}{1-\gamma}} p_k^{\frac{1}{1-\gamma}} A_k^{\frac{1}{1-\gamma}} \left(\frac{1}{w_k} \right)^{\frac{\gamma}{1-\gamma}} N_k^{\frac{\phi}{1-\gamma}} (z)^{\frac{\phi+\gamma-1}{1-\gamma}} = \Lambda_k (z)^{\frac{\phi+\gamma-1}{1-\gamma}}.
\end{aligned}$$

Thus, we have

$$\begin{aligned}
0 &= z^{\frac{\phi+\gamma-1}{1-\gamma}} \sum_k (1-\tau_k) \Lambda_k - (1-\tau_i) p_i - \sum_k \Lambda_k (z)^{\frac{\phi+\gamma-1}{1-\gamma}} (\tau_i - \tau_k) \\
z &= \left(\frac{\sum_k (1-\tau_k) \Lambda_k}{(1-\tau_i) p_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}}.
\end{aligned}$$

Hence we obtain 74. Now, for the profit shifting case, we can rewrite 70 as

$$\begin{aligned}
0 &= z^{\frac{\phi+\gamma-1}{1-\gamma}} \sum_k (1-\tau_k) \Lambda_k - (1-\tau_i) p_i - \sum_k \vartheta_k(z) (\tau_i - \tau_k) - \sum_k \vartheta_k(z) (1-\tau_i) \mathcal{C}(\lambda) + \lambda \sum_k \vartheta_k(z) [(1-\varphi)(\tau_i - \tau_{i^*})] \\
&= z^{\frac{\phi+\gamma-1}{1-\gamma}} \sum_k (1-\tau_k) \Lambda_k - (1-\tau_i) p_i - z^{\frac{\phi+\gamma-1}{1-\gamma}} \sum_k \Lambda_k [(\tau_i - \tau_k) + \lambda (1-\varphi)(\tau_i - \tau_{i^*}) - (1-\tau_i) \mathcal{C}(\lambda)],
\end{aligned}$$

and thus

$$\begin{aligned}
z &= \left(\frac{\sum_k \Lambda_k (1-\tau_i) \left[(1 - \mathcal{C}(\lambda)) + \frac{\lambda(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i} \right]}{(1-\tau_i) p_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}} \\
&= z^{TP} \left((1 - \mathcal{C}(\lambda)) + \frac{\lambda (1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}},
\end{aligned}$$

thus we have 75. ■

Now, we move towards proving Lemma 1.

Proof of Proposition 1. Note we have derived the formulas for $z^{FT}, z^{TP}$and z^{PS} and we have the following formulas for λ and $\mathcal{C}(\lambda)$:

$$\lambda = 1 - \exp\left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right), \tag{76}$$

$$\mathcal{C}(\lambda) = (\lambda - (\lambda - 1)\log(1-\lambda)), \tag{77}$$

where

$$\tau_{i^*} \equiv \min\{\tau_1, ..., \tau_K\}.$$

Start with showing 1. Let

$$\tau_i \equiv \max\{\tau_1, ..., \tau_K\},$$

then

$$1 - \tau_i < 1 - \tau_k \quad \forall k,$$
$$\frac{1-\tau_i}{1-\tau_i} < \frac{1-\tau_k}{1-\tau_i} \quad \forall k.$$

Thus

$$z^{FT} = \left(\frac{\sum_k (1-\tau_k)\Lambda_k}{(1-\tau_i)p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}} = \left(\frac{1}{p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}} \left(\sum_k \frac{(1-\tau_k)}{(1-\tau_i)}\Lambda_k\right)^{\frac{1-\gamma}{1-\phi-\gamma}}$$
$$> \left(\frac{1}{p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}} \left(\sum_k \frac{(1-\tau_i)}{(1-\tau_i)}\Lambda_k\right)^{\frac{1-\gamma}{1-\phi-\gamma}} = z^{TP},$$

thus we have

$$z^{TP} < z^{FT}.$$

Now, towards showing 2. Start with ($\Leftarrow$) direction, and let $0 < \varphi < 1$. Then, by 76 we have $0 < \lambda < 1$. Take any $\lambda \in (0,1)$ and notice that $z^{PS} > z^{TP}$ iff

$$\mathcal{C}(\lambda) < \frac{\lambda(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}. \tag{78}$$

Note that $\forall x > 0$ we have

$$x < -\ln(1-x)$$
$$-x > \ln(1-x)$$
$$\exp(-x) > -x + 1$$
$$1 - \exp(-x) < x$$
$$\left(\frac{1}{1-\exp(-x)}\right) > \frac{1}{x}$$
$$\left(1 + \frac{\exp(-x)}{1-\exp(-x)}\right) > \frac{1}{x}.$$

Now, set

$$x \equiv \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{(1-\tau_i)},$$

which implies

$$\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{(1-\tau_i)} \left[1 + \frac{\left(\exp\left(-\frac{a(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)\right)}{1-\exp\left(-\frac{a(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)} \right] > 1$$

$$\frac{\left(\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)\right)}{1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)} \left(\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{(1-\tau_i)}\right) > 1 - \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{(1-\tau_i)}$$

$$\frac{\left(-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)\right)}{1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)} \left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right) > 1 - \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{(1-\tau_i)}$$

$$\frac{\left(1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)-1\right)}{1-\exp\left(-\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)} \log\left(1-1+\exp\left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right)\right) > 1 - \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{(1-\tau_i)},$$

which using 76 can be written as

$$\frac{(\lambda-1)}{\lambda}\log(1-\lambda) > 1 - \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{(1-\tau_i)} > 0,$$

which through the series of iff inequalities can be transformed as follows

$$1 - \frac{(\lambda-1)}{\lambda}\log(1-\lambda) < \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}$$

$$\frac{(\lambda - (\lambda-1)\log(1-\lambda))}{\lambda} < \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}$$

$$\frac{\mathcal{C}(\lambda)}{\lambda} < \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}$$

$$\mathcal{C}(\lambda) < \frac{\lambda(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}.$$

This proves 78 and hence establishes $z^{PS} > z^{TP}$. Given that all the inequalities are iffs the reverse argument holds immediately. To show 3. and 4. notice from 76 first, that

$$\frac{\partial \lambda}{\partial \varphi} < 0.$$

Now, we want to show

$$\frac{\partial z^{PS}}{\partial \varphi} = z^{TP}\left((1-\mathcal{C}(\lambda)) + \frac{\lambda(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right)^{\frac{1-\gamma}{1-\phi+\gamma}-1} \times$$
$$\left(\left(-\mathcal{C}'(\lambda)\frac{\partial \lambda}{\partial \varphi}\right) + \frac{\partial \lambda}{\partial \varphi}\left[\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right] - \lambda\frac{(\tau_i - \tau_{i^*})}{1-\tau_i}\right)$$
$$= z^{PS}\left((1-\mathcal{C}(\lambda)) + \frac{\lambda(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right)^{-1}\left(\frac{\partial \lambda}{\partial \varphi}\left[\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i} - \mathcal{C}'(\lambda)\right] - \lambda\frac{(\tau_i - \tau_{i^*})}{1-\tau_i}\right) < 0.$$

This is negative if

$$\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}-\mathcal{C}'(\lambda)\leq 0,$$

and it holds with equality, since it is the condition equalizing marginal cost with marginal benefit of profit shifting λ. Thus, we get

$$\frac{\partial z^{PS}}{\partial \varphi}=z^{PS}\left((1-\mathcal{C}(\lambda))+\frac{\lambda(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)^{-1}\left(-\lambda\frac{(\tau_i-\tau_{i^*})}{1-\tau_i}\right)<0,$$

which proves 3. Notice, that proof for 4. follows analogously. Now towards deriving the elasticity

$$\begin{aligned}
\varepsilon^z_{\tau_{i^*}} &= \frac{1-\gamma}{1-\phi+\gamma}\left((1-\mathcal{C}(\lambda))+\frac{\lambda(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right)^{-1}\left(-\frac{\tau_{i^*}\lambda(1-\varphi)}{1-\tau_i}\right)\\
&= \frac{1-\gamma}{1-\phi+\gamma}\frac{-\tau_{i^*}\lambda(1-\varphi)}{\lambda(1-\varphi)(\tau_i-\tau_{i^*})\left[1+\frac{(1-\tau_i)(1-\mathcal{C}(\lambda))}{(\tau_i-\tau_{i^*})\lambda(1-\varphi)}\right]}\\
&= \frac{1-\gamma}{1-\phi+\gamma}\left(\frac{-\tau_{i^*}}{\tau_i-\tau_{i^*}}\right)\frac{1}{\left[1+\frac{1-\mathcal{C}(\lambda)}{\mathcal{C}'(\lambda)}\right]}<0.
\end{aligned}$$

■

A.1.1 Proofs under Alternative Assumption

Here, we assume that MNEs internalize the effect of changing z on the licensing fee $\vartheta_k(z)$ and solve for optimal z under different scenarios (FT, TP, and PS). We then prove Proposition 1 under this assumption. Note that the optimal shifting share λ will not be changed as it is solved independently from z so Lemma 1 holds automatically. Let's first solve for optimal z under this assumption.

Lemma 4 *The allocations of intangible capital are as follows:*

$$z^{FT}=\left(\frac{\sum_k(1-\tau_k)\Lambda_k}{(1-\tau_i)p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}, \tag{79}$$

$$z^{TP}=\left[\frac{\sum_k(1-\tau_k)\Lambda_k-\frac{\phi}{1-\gamma}\sum_k(\tau_i-\tau_k)\Lambda_k}{(1-\tau_i)p_i}\right]^{\frac{1-\gamma}{1-\phi-\gamma}}, \tag{80}$$

$$z^{PS}=\left[\frac{-\frac{\phi}{1-\gamma}\mathcal{C}(\lambda)\sum_k\Lambda_k}{p_i}+\frac{\sum_k(1-\tau_k)\Lambda_k-\frac{\phi}{1-\gamma}\sum_k(\tau_i-\tau_k)\Lambda_k+\lambda\frac{\phi}{1-\gamma}(\tau_i-\tau_{i^*})(1-\varphi)\sum_k\Lambda_k}{(1-\tau_i)p_i}\right]^{\frac{1-\gamma}{1-\phi-\gamma}}. \tag{81}$$

Proof of Lemma 4. Starting from the profit maximization problem of an MNE:

$$\max_{z,\lambda,\{l_i\}_{i=1}^I} (1-\tau_i)\left(p_i\left(A_i\left(N_iz\right)^{\phi}l_i^{\gamma}\right)-w_il_i-p_iz+z\left[\varphi\lambda\sum_k\vartheta_k\left(z\right)-\lambda\vartheta_i\left(z\right)+\left(1-\lambda\right)\sum_{k\neq i}\vartheta_k\left(z\right)-\sum_k\vartheta_k\left(z\right)\mathcal{C}\left(\lambda\right)\right]\right)$$
$$+\left(1-\tau_{i^*}\right)\left(p_{i^*}\left(A_{i^*}\left(N_{i^*}z\right)^{\phi}l_{i^*}^{\gamma}\right)-w_kl_{i^*}+z\left[\lambda\sum_{k\neq i^*}\vartheta_k\left(z\right)-\left(1-\lambda\right)\vartheta_{i^*}\left(z\right)-\varphi\lambda\sum_k\vartheta_k\left(z\right)\right]\right)$$
$$+\left(1-\tau_k\right)\sum_{k\neq i,i^*}\left(p_k\left(A_k\left(N_kz\right)^{\phi}l_k^{\gamma}\right)-w_kl_k-\vartheta_k\left(z\right)z\right). \tag{82}$$

With the derivative of $\vartheta_k(z)$ with respect to z taken, the FOC with respect to z is then:

$$0=\sum_k\left(1-\tau_k\right)\phi N_kp_kA_k\left(N_kz\right)^{\phi-1}l_k^{\gamma}-\left(1-\tau_i\right)p_i$$
$$-\left[\sum_{k\neq i,i^*}\left(1-\tau_k\right)\vartheta_k\left(z\right)-\sum\left(1-\tau_i\right)\vartheta_k\left(z\right)+\left(1-\tau_{i^*}\right)\vartheta_{i^*}\left(z\right)\right]-\left(1-\tau_i\right)\sum_k\vartheta_k\left(z\right)\mathcal{C}\left(\lambda\right)$$
$$-z\left[\sum_{k\neq i,i^*}\left(1-\tau_k\right)\vartheta_k'\left(z\right)-\sum_{k\neq i}\left(1-\tau_i\right)\vartheta_k'\left(z\right)+\left(1-\tau_{i^*}\right)\vartheta_{i^*}'\left(z\right)\right]-\left(1-\tau_i\right)z\sum_k\vartheta_k'\left(z\right)\mathcal{C}\left(\lambda\right)$$
$$+\lambda\Bigg[\left(1-\tau_i\right)\varphi\sum_k\vartheta_k\left(z\right)-\left(1-\tau_i\right)\vartheta_i\left(z\right)-\left(1-\tau_i\right)\sum_{k\neq i}\vartheta_k\left(z\right)+\left(1-\tau_{i^*}\right)\sum_{k\neq i^*}\vartheta_k\left(z\right)+$$
$$\left(1-\tau_{i^*}\right)\vartheta_{i^*}\left(z\right)-\left(1-\tau_{i^*}\right)\varphi\sum_k\vartheta_k\left(z\right)\Bigg]$$
$$+\lambda z\Bigg[\left(1-\tau_i\right)\varphi\sum_k\vartheta_k'\left(z\right)-\left(1-\tau_i\right)\vartheta_i'\left(z\right)-\left(1-\tau_i\right)\sum_{k\neq i}\vartheta_k'\left(z\right)+\left(1-\tau_{i^*}\right)\sum_{k\neq i^*}\vartheta_k'\left(z\right)+$$
$$\left(1-\tau_{i^*}\right)\vartheta_{i^*}'\left(z\right)-\left(1-\tau_{i^*}\right)\varphi\sum_k\vartheta_k'\left(z\right)\Bigg]. \tag{83}$$

Now plug the optimal labor $l_i=\left(\gamma p_iA_i\left(N_iz\right)^{\phi}w_i^{-1}\right)^{\frac{1}{1-\gamma}}$; then we can derive as before

$$\vartheta_k\left(z\right)=\Lambda_k\cdot z^{\frac{\phi+\gamma-1}{1-\gamma}},$$

and

$$\vartheta_k'\left(z\right)=\frac{\phi+\gamma-1}{1-\gamma}\Lambda_k\cdot z^{\frac{\phi+\gamma-1}{1-\gamma}-1}.$$

Plugging them back to the FOC of z and further simplifying:

$$
\begin{aligned}
(1-\tau_i)\, p_i = & \sum_k (1-\tau_k)\,\Lambda_k z^{\frac{\phi+\gamma-1}{1-\gamma}} \\
& - \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} \left[\sum_{k\neq i,i^*} (1-\tau_k)\,\Lambda_k - (1-\tau_i) \sum_{k\neq i} \Lambda_k + (1-\tau_{i^*})\,\Lambda_{i^*} \right] \\
& - \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} \mathcal{C}(\lambda)\,(1-\tau_i) \sum_k \Lambda_k \\
& + \lambda \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} \Bigg[(1-\tau_i)\,\varphi \sum_k \Lambda_k - (1-\tau_i)\,\Lambda_k - (1-\tau_i) \sum_{k\neq i} \Lambda_k + \\
& \quad (1-\tau_{i^*}) \sum_{k\neq i^*} \Lambda_k + (1-\tau_{i^*})\,\Lambda_k - (1-\tau_{i^*})\,\varphi \sum_k \Lambda_k \Bigg] \\
= & \sum_k (1-\tau_k)\,\Lambda_k z^{\frac{\phi+\gamma-1}{1-\gamma}} - \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} \sum_k (\tau_i-\tau_k)\,\Lambda_k - \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} \mathcal{C}(\lambda)\,(1-\tau_i) \sum_k \Lambda_k \\
& - \lambda \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} (\tau_{i^*}-\tau_i)\,(\varphi-1) \sum_k \Lambda_k .
\end{aligned}
$$

The case of free transfer requires $\vartheta_k(z) = 0$ and the solution of z is the same as before

$$
z = \left(\frac{\sum_k (1-\tau_k)\,\Lambda_k}{(1-\tau_i)\, p_i} \right)^{\frac{1-\gamma}{1-\phi-\gamma}} . \tag{84}
$$

The case of transfer pricing requires $\lambda = 0$, then

$$
\begin{aligned}
(1-\tau_i)\, p_i = & \sum_k (1-\tau_k)\,\Lambda_k z^{\frac{\phi+\gamma-1}{1-\gamma}} \\
& - \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} \left[\sum_{k\neq i,i^*} (1-\tau_k)\,\Lambda_k - (1-\tau_i) \sum_{k\neq i} \Lambda_k + (1-\tau_{i^*})\,\Lambda_{i^*} \right] \\
= & \sum_k (1-\tau_k)\,\Lambda_k z^{\frac{\phi+\gamma-1}{1-\gamma}} - \frac{\phi}{1-\gamma} z^{\frac{\phi+\gamma-1}{1-\gamma}} \sum_k (\tau_i-\tau_k)\,\Lambda_k .
\end{aligned}
$$

We can solve for z as:

$$
z = \left[\frac{\sum_k (1-\tau_k)\,\Lambda_k - \frac{\phi}{1-\gamma} \sum_k (\tau_i-\tau_k)\,\Lambda_k}{(1-\tau_i)\, p_i} \right]^{\frac{1-\gamma}{1-\phi-\gamma}} , \tag{85}
$$

thus, we obtain 80. Now turn to the *PS* case, we can solve for z as:

$$
z = \left[\frac{-\frac{\phi}{1-\gamma}\mathcal{C}(\lambda) \sum_k \Lambda_k}{p_i} + \frac{\sum_k (1-\tau_k)\,\Lambda_k - \frac{\phi}{1-\gamma} \sum_k (\tau_i-\tau_k)\,\Lambda_k + \lambda \frac{\phi}{1-\gamma} (\tau_i-\tau_{i^*})\,(1-\varphi) \sum_k \Lambda_k}{(1-\tau_i)\, p_i} \right]^{\frac{1-\gamma}{1-\phi-\gamma}} , \tag{86}
$$

thus, we obtain 81. ■

With these optimal intangible allocations derived, we now prove Lemma 1 under the alternative assumption. Note that we will not obtain the same elasticity results of z^{PS} with respect to φ and τ_{i^*} but rather show the comparative statics results hold, namely z^{PS} is decreasing in both φ and τ_{i^*}.

Proof of Proposition 1 under alternative assumption. Start from 1, it's obvious that when $\tau_i \equiv \max\{\tau_1, ..., \tau_K\}$:

$$z^{FT} = \left(\frac{\sum_k (1-\tau_k)\Lambda_k}{(1-\tau_i)p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}} > \left[\frac{\sum_k (1-\tau_k)\Lambda_k - \frac{\phi}{1-\gamma}\sum_k (\tau_i - \tau_k)\Lambda_k}{(1-\tau_i)p_i}\right]^{\frac{1-\gamma}{1-\phi-\gamma}} = z^{TP}.$$

Now, towards showing 2. Start with ($\Leftarrow$) direction, and let $0 < \varphi < 1$. Then, by 76 we have $0 < \lambda < 1$. Take any $\lambda \in (0,1)$ and notice that $z^{PS} > z^{TP}$ iff

$$\mathcal{C}(\lambda) < \frac{\lambda(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i},$$

which has already been proven above. Given that all the inequalities are iffs the reverse argument holds immediately. Hence, we prove 2. To show 3. and 4. Notice from 76 first that $\frac{\partial\lambda}{\partial\varphi} < 0$. Now, we want to show

$$\begin{aligned}\frac{\partial z^{PS}}{\partial\varphi} &= \frac{1-\gamma}{1-\phi-\gamma}\left(z^{PS}\right)^{\frac{\phi}{1-\gamma}}\frac{\sum_k \Lambda_k}{p_i}\frac{\phi}{1-\gamma}\left(\left(-\mathcal{C}'(\lambda)\frac{\partial\lambda}{\partial\varphi}\right) + \frac{\partial\lambda}{\partial\varphi}\left[\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i}\right] - \lambda\frac{(\tau_i-\tau_{i^*})}{1-\tau_i}\right)\\ &= \frac{1-\gamma}{1-\phi-\gamma}\left(z^{PS}\right)^{\frac{\phi}{1-\gamma}}\frac{\sum_k \Lambda_k}{p_i}\frac{\phi}{1-\gamma}\left(\frac{\partial\lambda}{\partial\varphi}\left[\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i} - \mathcal{C}'(\lambda)\right] - \lambda\frac{(\tau_i-\tau_{i^*})}{1-\tau_i}\right) < 0,\end{aligned}$$

which is true if

$$\frac{(1-\varphi)(\tau_i-\tau_{i^*})}{1-\tau_i} - \mathcal{C}'(\lambda) \leq 0.$$

And it holds with equality, since it is the condition equalizing marginal cost with marginal benefit of profit shifting λ. Thus, we get

$$\frac{\partial z^{PS}}{\partial\varphi} = \frac{1-\gamma}{1-\phi-\gamma}\left(z^{PS}\right)^{\frac{\phi}{1-\gamma}}\frac{\sum_k \Lambda_k}{p_i}\frac{\phi}{1-\gamma}\left(-\lambda\frac{(\tau_i-\tau_{i^*})}{1-\tau_i}\right) < 0,$$

which proves 3. Notice, that proof for 4. follows analogously when Λ_{i^*} is sufficiently small, which is the empirically relevant case for us.

$$\begin{aligned}\frac{\partial z^{PS}}{\partial\tau_{i^*}} &= \frac{1-\gamma}{1-\phi-\gamma}\left(z^{PS}\right)^{\frac{\phi}{1-\gamma}}\left(\frac{-\Lambda_{i^*} + \frac{\phi}{1-\gamma}\Lambda_{i^*} - \lambda\frac{\phi}{1-\gamma}(1-\varphi)\sum_k \Lambda_k}{(1-\tau_i)p_i}\right)\\ &= \frac{1-\gamma}{1-\phi-\gamma}\left(z^{PS}\right)^{\frac{\phi}{1-\gamma}}\left(\frac{\frac{\phi+\gamma-1}{1-\gamma}\Lambda_{i^*} - \lambda\frac{\phi}{1-\gamma}(1-\varphi)\sum_k \Lambda_k}{(1-\tau_i)p_i}\right) < 0.\end{aligned}$$

■

A.2 Allocation of Profit Rule

The proof of Lemma 2 is provided below. First to simplify notation we denote

$$\hat{\tau}_i(\theta) = (1-\theta)\tau_i + \theta \sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k}.$$

Proof of Lemma 2. Start with derivitation of the optimal λ in the case of profit reallocation. Recall that the profit maximization problem of the MNE is

$$\begin{aligned}
\max_{z,\lambda,\{l_i\}_{i=1}^I} \; & (1-\tau_i(1-\theta))\Big(p_i\left(A_i(N_i z)^\phi l_i^\gamma\right) - w_i l_i - p_i z \\
& + z\left[\varphi\lambda\sum_k \vartheta_k(z) - \lambda\vartheta_i(z) + (1-\lambda)\sum_{k\neq i}\vartheta_k(z) - C(\lambda)\sum_k \vartheta_k(z)\right]\Big) \\
& + (1-\tau_{i^*}(1-\theta))\Big(p_{i^*}\left(A_{i^*}(N_{i^*}z)^\phi l_{i^*}^\gamma\right) - w_k l_{i^*} \\
& + z\left[\lambda\sum_{k\neq i^*}\vartheta_k(z) - (1-\lambda)\vartheta_{i^*}(z) - \varphi\lambda\sum_k \vartheta_k(z)\right]\Big) \\
& + (1-\tau_k(1-\theta))\sum_{k\neq i,i^*}\Big(p_k\left(A_k(N_k z)^\phi l_k^\gamma\right) - w_k l_k - \vartheta_k(z)\,z\Big) \\
& - (1-(1-\theta))\sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k}\cdot\left[\sum_k\Big(p_k\left(A_k(N_k z)^\phi l_k^\gamma\right) - w_k l_k\Big) - p_i z - C(\lambda)\sum_k \vartheta_k(z)\right].
\end{aligned}$$

Take the derivative with respect to λ:

$$\begin{aligned}
\lambda : 0 = & (\varphi-1)\sum_k \vartheta_k(z)\left[(1-\tau_i(1-\theta)) - (1-\tau_{i^*}(1-\theta))\right] \\
& - (1-\tau_i(1-\theta))\sum_k \vartheta_k(z)\,\mathcal{C}'(\lambda) + \theta\sum \vartheta_k(z)\,\mathcal{C}'(\lambda)\cdot\sum_i \tau_i\cdot\frac{p_i y_i}{\sum_k p_k y_k},
\end{aligned}$$

and rearranging we get

$$\mathcal{C}'(\lambda)\cdot\left[(1-\tau_i(1-\theta)) - \theta\sum_i \tau_i\cdot\frac{p_i y_i}{\sum_k p_k y_k}\right] = (1-\theta)(\varphi-1)(\tau_{i^*}-\tau_i),$$

and we can derive:

$$\begin{aligned}
\mathcal{C}'(\lambda) &= (1-\varphi)\frac{(1-\theta)(\tau_i-\tau_{i^*})}{(1-\tau_i(1-\theta)) - \theta\sum_i \tau_i\cdot\frac{p_i y_i}{\sum_k p_k y_k}} \\
\lambda &= (\mathcal{C}')^{-1}\left[\frac{(1-\varphi)(1-\theta)(\tau_i-\tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right].
\end{aligned}$$

Parametrizing $\mathcal{C}(\lambda) = (\lambda + (1-\lambda)\log(1-\lambda))$, we can solve for λ as:

$$\hat{\lambda} = 1 - \exp\left(-\frac{(1-\varphi)(1-\theta)(\tau_i-\tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right). \tag{87}$$

Now, compare equation 87 with it's counterpart under the current tax regime, which is given by

$$\lambda = 1 - \exp\left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right).$$

Towards proving 1., pick any $0 < \theta \leq 1$. Then the following sequence of inequalities holds:

$$\begin{aligned}
1 - \exp\left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right) &> 1 - \exp\left(-\frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right) \\
\exp\left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right) &< \exp\left(-\frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right) \\
\left(-\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}\right) &< \left(-\frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right) \\
\frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i} &> \frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)} \\
\frac{1}{1-\tau_i} &> \frac{1-\theta}{1-\hat{\tau}_i(\theta)} \\
1 &> \frac{1-(1-\theta)\tau_i - \theta}{1-(1-\theta)\tau_i - \theta\sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k}} \\
1-(1-\theta)\tau_i - \theta\sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k} &> 1-(1-\theta)\tau_i - \theta \\
-\theta\sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k} &> -\theta \\
\sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k} &< 1.
\end{aligned}$$

The last inequality holds, since $\tau_k < 1 \ \forall k$ and all sales shares are by construction less than one. This proves that $\hat{\lambda} < \lambda$. Now, towards showing 2, inspect how θ affects λ, i.e.

$$\begin{aligned}
\frac{\partial \lambda(\theta)}{\partial \theta} &= -\exp\left(-\frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right) \cdot \\
&\quad (-1)\left(\frac{-(1-\varphi)(\tau_i - \tau_{i^*})\left[1-\hat{\tau}_i(\theta)\right] + (1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})\left[\tau_i - \sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k}\right]}{\left[1-\hat{\tau}_i(\theta)\right]^2}\right) \\
&= -(1-\varphi)(\tau_i - \tau_{i^*})\exp\left(-\frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right)\frac{1-\sum_i \tau_i \cdot \frac{p_i y_i}{\sum_k p_k y_k}}{\left[1-\hat{\tau}_i(\theta)\right]^2} < 0,
\end{aligned}$$

and the elasticity is given

$$\begin{aligned}
\varepsilon_{\theta}^{\lambda} = \frac{\partial\lambda\left(\theta\right)}{\partial\theta}\frac{\theta}{\lambda} &= -\left(1-\varphi\right)\left(\tau_i-\tau_{i^*}\right)\exp\left(-\frac{\left(1-\varphi\right)\left(1-\theta\right)\left(\tau_i-\tau_{i^*}\right)}{1-\hat{\tau}_i(\theta)}\right)\frac{1-\sum_i\tau_i\cdot\frac{p_iy_i}{\sum_k p_ky_k}}{\left[1-\hat{\tau}_i(\theta)\right]^2}\frac{\theta}{\lambda}\\
&= -\left(\frac{1-\lambda}{\lambda}\right)\left(1-\varphi\right)\left(\tau_i-\tau_{i^*}\right)\frac{\theta\left(1-\sum_i\tau_i\cdot\frac{p_iy_i}{\sum_k p_ky_k}\right)}{\left[1-\hat{\tau}_i(\theta)\right]^2}\\
&= -\left(\frac{1-\lambda}{\lambda}\right)\left(1-\varphi\right)\left(\tau_i-\tau_{i^*}\right)\frac{\theta\left(1-\sum_i\tau_i\cdot\frac{p_iy_i}{\sum_k p_ky_k}\right)}{\left[1-\hat{\tau}_i(\theta)\right]^2}\\
&= -\left(\frac{1-\lambda}{\lambda}\right)\mathcal{C}'\left(\lambda\right)\frac{\theta}{1-\theta}\frac{\left(1-\sum_i\tau_i\cdot\frac{p_iy_i}{\sum_k p_ky_k}\right)}{\left[1-\hat{\tau}_i(\theta)\right]}<0,
\end{aligned}$$

where in the last equality we used the first order condition of the profit function with respect to λ. Hence, we have established 2. Now, inspect how τ_{i^*} affects λ to prove 3. First, compute the relevant partial derivative

$$\begin{aligned}
\frac{\partial\lambda\left(\theta\right)}{\partial\tau_{i^*}} &= -\exp\left(-\frac{\left(1-\varphi\right)\left(1-\theta\right)\left(\tau_i-\tau_{i^*}\right)}{1-\hat{\tau}_i(\theta)}\right)\cdot\\
&\left(-1\right)\left(\frac{-\left(1-\varphi\right)\left(1-\theta\right)\left[1-\hat{\tau}_i(\theta)\right]+\left(1-\varphi\right)\left(1-\theta\right)\left(\tau_i-\tau_{i^*}\right)\left[-\frac{p_{i^*}y_{i^*}}{\sum_k p_ky_k}\right]}{\left[1-\hat{\tau}_i(\theta)\right]^2}\right)<0,
\end{aligned}$$

and hence the elasticity

$$\begin{aligned}
\varepsilon_{\tau_{i^*}}^{\hat{\lambda}} &= -\exp\left(-\frac{\left(1-\varphi\right)\left(1-\theta\right)\left(\tau_i-\tau_{i^*}\right)}{1-\hat{\tau}_i(\theta)}\right)\cdot\\
&\left(-1\right)\left(\frac{-\left(1-\varphi\right)\left(1-\theta\right)\left[1-\hat{\tau}_i(\theta)\right]+\left(1-\varphi\right)\left(1-\theta\right)\left(\tau_i-\tau_{i^*}\right)\left[-\frac{p_{i^*}y_{i^*}}{\sum_k p_ky_k}\right]}{\left[1-\hat{\tau}_i(\theta)\right]^2}\right)\frac{\tau_{i^*}}{\lambda}\\
&= \varepsilon_{\tau_{i^*}}^{\lambda}\left[\left(1-\tau_i\right)\left(1-\theta\right)\left(\frac{\left[1-\hat{\tau}_i(\theta)\right]+\left(\tau_i-\tau_{i^*}\right)\left(\frac{p_{i^*}y_{i^*}}{\sum_k p_ky_k}\right)}{\left[1-\hat{\tau}_i(\theta)\right]^2}\right)\right].
\end{aligned}$$

Suppose that the size of tax-haven is negligible i.e. $p_{i^*}y_{i^*}\approx 0$, then we have

$$\begin{aligned}
\varepsilon_{\tau_{i^*}}^{\hat{\lambda}} &= \varepsilon_{\tau_{i^*}}^{\lambda}\left(1-\tau_i\right)\left(1-\theta\right)\left(\frac{\left[1-\hat{\tau}_i(\theta)\right]}{\left[1-\hat{\tau}_i(\theta)\right]^2}\right)\\
&= \varepsilon_{\tau_{i^*}}^{\lambda}\left(\frac{1-\tau_i\left(1-\theta\right)-\theta}{\left[1-\hat{\tau}_i(\theta)\right]}\right),
\end{aligned}$$

and note that

$$\begin{aligned}
&1-\left(1-\theta\right)\tau_i-\theta<1-\left(1-\theta\right)\tau_i-\theta\sum_i\tau_i\cdot\frac{p_iy_i}{\sum_k p_ky_k}=1-\hat{\tau}_i(\theta)\\
&\frac{1-\tau_i\left(1-\theta\right)-\theta}{\left[1-\hat{\tau}_i(\theta)\right]}<1,
\end{aligned}$$

which implies immediately

$$\left|\varepsilon_{\tau_{i^*}}^{\hat{\lambda}}\right| < \left|\varepsilon_{\tau_{i^*}}^{\lambda}\right|.$$

■

We now move on to prove Proposition 2. We first derive the formulas for allocation of the intangible capital under the profit allocation rule. The following lemma summarizes them.

Lemma 5 *The allocations of intangible capital and share of shifted intangible capital under the profit allocation rule are as follows:*

$$\begin{aligned}\hat{z}^{FT} &= \left(\frac{\sum_k (1-\tau_k)\Lambda_k}{[1-\hat{\tau}_i(\theta)]\, p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}},\\ \hat{z}^{TP} &= \left(\frac{\sum_k \Lambda_k}{p_i}\right)^{\frac{1-\gamma}{1-\phi+\gamma}},\\ \hat{z}^{PS} &= \hat{z}^{TP}\left(1 - C(\lambda) + \frac{(1-\theta)\,\lambda\,(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}.\end{aligned}$$

Proof of Lemma 5. The proof follows the same procedure as Lemma 3 . ■

With Lemma 5 at hand, we are in a position to prove Lemma 2.

Proof of Proposition 2. We begin with proving 1. The proof relies on the following sequence of iff inequalities:

$$\begin{aligned}\hat{z}^{PS} &< z^{PS}\\ \hat{z}^{TP}\left(1 - C\left(\hat{\lambda}\right) + \frac{(1-\theta)\,\hat{\lambda}\,(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right)^{\frac{1-\gamma}{1-\phi-\gamma}} &< z^{TP}\left(1 - \mathcal{C}\left(\lambda\right) + \frac{\lambda\,(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}\\ \left(1 - C\left(\hat{\lambda}\right) + \hat{\lambda}\left[\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right]\frac{(1-\tau_i)\,(1-\theta)}{1-\hat{\tau}_i(\theta)}\right) &< \left(1 - \mathcal{C}\left(\lambda\right) + \frac{\lambda\,(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right)\\ \lambda\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i} - \hat{\lambda}\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})\,(1-\theta)}{1-\hat{\tau}_i(\theta)} &> \mathcal{C}\left(\lambda\right) - \mathcal{C}\left(\hat{\lambda}\right),\end{aligned}$$

where we have

$$\begin{aligned}\mathcal{C}\left(\lambda\right) &\equiv \lambda + (1-\lambda)\log\left(1-\lambda\right)\\ \hat{\lambda} &= 1 - \exp\left(-\frac{(1-\varphi)\,(1-\theta)\,(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)}\right)\\ \lambda &= 1 - \exp\left(-\frac{(1-\varphi)\,(\tau_i - \tau_{i^*})}{1-\tau_i}\right).\end{aligned}$$

To simplify notation, let's denote:

$$\hat{A} = \frac{(1-\varphi)(1-\theta)(\tau_i - \tau_{i^*})}{1-\hat{\tau}_i(\theta)},$$
$$A = \frac{(1-\varphi)(\tau_i - \tau_{i^*})}{1-\tau_i}.$$

Plugging back to the inequality we want to show, we have

$$\begin{aligned}
1-\exp(-A)+\exp(-A)(-A)-1+\exp\left(-\hat{A}\right)-\exp\left(-\hat{A}\right)\left(-\hat{A}\right) &< (1-\exp(-A))A-\left(1-\exp\left(-\hat{A}\right)\right)\hat{A} \\
-\exp(-A)+\exp(-A)(-A)+\exp\left(-\hat{A}\right)-\exp\left(-\hat{A}\right)\left(-\hat{A}\right) &< A-\hat{A}-\exp(-A)A+\exp\left(-\hat{A}\right)\hat{A} \\
-\exp(-A)+\exp\left(-\hat{A}\right) &< A-\hat{A} \\
\hat{A}+\exp\left(-\hat{A}\right) &< A+\exp(-A).
\end{aligned}$$

We have shown that $0<\hat{A}<A,\ \theta>0$, thus proving the inequality above amounts to prove that function $f(x)=x+\exp(-x)$ is monotonically increasing when $x>0$. Taking its derivative we get:

$$f'(x) = 1-\exp(-x) > 0,\ x>0.$$

To prove 2., we start with the appropriate partial derivative with respect to θ, i.e.

$$\frac{\partial \widehat{z}^{PS}}{\partial\theta} = \frac{\partial \hat{z}^{TP}}{\partial\theta} + \frac{\partial\left(1-C(\lambda)+\frac{(1-\theta)\lambda(1-\varphi)(\tau_i-\tau_{i^*})}{(1-\hat{\tau}_i(\theta))}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}}{\partial\theta}.$$

We know that $\frac{\partial \widehat{z}^{PS}}{\partial\theta}=0$. Then the sign of $\frac{\partial \widehat{z}^{PS}}{\partial\theta}$ is determined by the second part, which has the same sign as:

$$\begin{aligned}
&\left[-C'(\lambda)\frac{\partial\lambda}{\partial\theta}+\frac{\left[\left(-\lambda+(1-\theta)\frac{\partial\lambda}{\partial\theta}\right)(1-\varphi)(\tau_i-\tau_{i^*})\right](1-\hat{\tau}_i(\theta))-\left(\left(\tau_i-\frac{\sum_k\Lambda_k\tau_k}{\sum_k\Lambda_k}\right)\right)(1-\theta)\lambda(1-\varphi)(\tau_i-\tau_{i^*})}{\left[(1-\hat{\tau}_i(\theta))\right]^2}\right] \\
&=\frac{\partial\lambda}{\partial\theta}\left[\frac{(1-\theta)(1-\varphi)(\tau_i-\tau_{i^*})}{(1-\hat{\tau}_i(\theta))}-C'(\lambda)\right]+ \\
&\frac{-\lambda(1-\varphi)(\tau_i-\tau_{i^*})(1-\hat{\tau}_i(\theta))-\left(\left(\tau_i-\frac{\sum_k\Lambda_k\tau_k}{\sum_k\Lambda_k}\right)\right)(1-\theta)\lambda(1-\varphi)(\tau_i-\tau_{i^*})}{\left[(1-\hat{\tau}_i(\theta))\right]^2},
\end{aligned}$$

and notice that the FOC w.r.t. λ is given by

$$\mathcal{C}'(\lambda) = (1-\varphi)\frac{(1-\theta)(\tau_i-\tau_{i^*})}{1-\hat{\tau}_i(\theta)}.$$

Thus to evaluate the sign, we need to sign the following

$$\left[\frac{-\lambda\left(1-\varphi\right)\left(\tau_{i}-\tau_{i^{*}}\right)\left(1-\hat{\tau}_{i}(\theta)\right)-\left(\left(\tau_{i}-\frac{\sum_{k}\Lambda_{k}\tau_{k}}{\sum_{k}\Lambda_{k}}\right)\right)\left(1-\theta\right)\lambda\left(1-\varphi\right)\left(\tau_{i}-\tau_{i^{*}}\right)}{\left[\left(1-\hat{\tau}_{i}(\theta)\right)\right]^{2}}\right]$$

$$=\left[\frac{\lambda\left(1-\varphi\right)\left(\tau_{i}-\tau_{i^{*}}\right)\left[-1+\left(1-\theta\right)\tau_{i}-\left(1-\theta\right)\tau_{i}+\frac{\sum_{k}\Lambda_{k}\tau_{k}}{\sum_{k}\Lambda_{k}}\right]}{\left[\left(1-\hat{\tau}_{i}(\theta)\right)\right]^{2}}\right]$$

$$=\left[\frac{\lambda\left(1-\varphi\right)\left(\tau_{i}-\tau_{i^{*}}\right)\left[\frac{\sum_{k}\Lambda_{k}\tau_{k}}{\sum_{k}\Lambda_{k}}-1\right]}{\left[\left(1-\hat{\tau}_{i}(\theta)\right)\right]^{2}}\right]<0,$$

thus we have established that

$$\frac{\partial\widehat{z}^{PS}}{\partial\theta}<0.$$

And the elasticity is

$$\varepsilon_{\theta}^{\hat{z}^{PS}}=\frac{\partial\widehat{z}^{PS}}{\partial\theta}\frac{\theta}{\widehat{z}^{PS}}=\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\widehat{z}^{PS}\frac{\theta}{\widehat{z}^{PS}}\left[\frac{\lambda\left(1-\varphi\right)\left(\tau_{i}-\tau_{i^{*}}\right)\left[\frac{\sum_{k}\Lambda_{k}\tau_{k}}{\sum_{k}\Lambda_{k}}-1\right]}{\left[\left(1-\hat{\tau}_{i}(\theta)\right)\right]^{2}}\right]\cdot$$

$$\frac{\left(1-\hat{\tau}_{i}(\theta)\right)}{\left(1-C(\lambda)\right)\left(1-\hat{\tau}_{i}(\theta)\right)+\left(1-\theta\right)\lambda\left(1-\varphi\right)\left(\tau_{i}-\tau_{i^{*}}\right)}$$

$$=\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\theta\left(\widehat{z}^{PS}\right)^{\frac{1-\phi-\gamma}{1-\gamma}}\left[\frac{\lambda\left(1-\varphi\right)\left(\tau_{i}-\tau_{i^{*}}\right)\left[\frac{\sum_{k}\Lambda_{k}\tau_{k}}{\sum_{k}\Lambda_{k}}-1\right]}{\left[\left(1-\hat{\tau}_{i}(\theta)\right)\right]^{2}}\right]<0.$$

Now, to show 3. consider $\hat{z}^{PS}$ with respect to $\tau_{i^{*}}$, which yields

$$\varepsilon_{\tau_{i^{*}}}^{\hat{z}^{PS}}=\frac{\partial\hat{z}^{PS}}{\partial\tau_{i^{*}}}\frac{\tau_{i^{*}}}{\hat{z}^{PS}},$$

where

$$\frac{\partial\hat{z}^{PS}}{\partial\tau_{i^{*}}}=\frac{\partial\hat{z}^{TP}}{\partial\tau_{i^{*}}}+\frac{\partial\left(1-C(\lambda)+\frac{(1-\theta)\lambda(1-\varphi)(\tau_{i}-\tau_{i^{*}})}{[1-\hat{\tau}_{i}(\theta)]}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}}{\partial\tau_{i^{*}}},$$

and $\frac{\partial\left(1-C(\lambda)+\frac{(1-\theta)\lambda(1-\varphi)(\tau_i-\tau_{i^*})}{[1-\hat{\tau}_i(\theta)]}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}}{\partial\tau_{i^*}}$ is given by

$$\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\hat{z}^{PS}\left(1-C(\lambda)+\frac{(1-\theta)\,\lambda\,(1-\varphi)\,(\tau_i-\tau_{i^*})}{[1-\hat{\tau}_i(\theta)]}\right)^{-1}\cdot$$
$$\left[\frac{\partial\hat{\lambda}}{\partial\tau_{i^*}}\left(\frac{(1-\varphi)\,(\tau_i-\tau_{i^*})\,(1-\theta)\,[1-\hat{\tau}_i(\theta)]}{[1-\hat{\tau}_i(\theta)]^2}-\mathcal{C}'\left(\hat{\lambda}\right)\right)+\right.$$
$$\left.\left(\frac{\theta\frac{\Lambda_{i^*}}{\sum_k\Lambda_k}\,(1-\theta)\,\lambda\,(1-\varphi)\,(\tau_i-\tau_{i^*})-(1-\theta)\,\hat{\lambda}\,(1-\varphi)\,[1-\hat{\tau}_i(\theta)]}{[1-\hat{\tau}_i(\theta)]^2}\right)\right]$$
$$=\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\hat{z}^{PS}\left(1-C(\lambda)+\frac{(1-\theta)\,\lambda\,(1-\varphi)\,(\tau_i-\tau_{i^*})}{[1-\hat{\tau}_i(\theta)]}\right)^{-1}\cdot$$
$$\left[\frac{\partial\hat{\lambda}}{\partial\tau_{i^*}}\left(\frac{(1-\varphi)\,(\tau_i-\tau_{i^*})\,(1-\theta)}{1-\left((1-\theta)\,\tau_i+\theta\sum_k\tau_k\frac{\Lambda_k}{\sum_k\Lambda_k}\right)}-\mathcal{C}'\left(\hat{\lambda}\right)\right)+\right.$$
$$\left.\left(\frac{\theta\frac{\Lambda_{i^*}}{\sum_k\Lambda_k}\,(1-\theta)\,\lambda\,(1-\varphi)\,(\tau_i-\tau_{i^*})-(1-\theta)\,\hat{\lambda}\,(1-\varphi)\,[1-\hat{\tau}_i(\theta)]}{[1-\hat{\tau}_i(\theta)]^2}\right)\right].$$

Notice that the FOC wrt to $\hat{\lambda}$ implies

$$\mathcal{C}'\left(\hat{\lambda}\right)=(1-\varphi)\,\frac{(1-\theta)\,(\tau_i-\tau_{i^*})}{1-\hat{\tau}_i(\theta)},$$

thus the elasticity becomes

$$\varepsilon_{\tau_{i^*}}^{\hat{z}^{PS}}=\frac{\tau_{i^*}}{\hat{z}^{PS}}\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\hat{z}^{PS}\left(1-C(\lambda)+\frac{(1-\theta)\,\lambda\,(1-\varphi)\,(\tau_i-\tau_{i^*})}{[1-\hat{\tau}_i(\theta)]}\right)^{-1}$$
$$\left[\left(\frac{\theta\frac{\Lambda_{i^*}}{\sum_k\Lambda_k}\,(1-\theta)\,\lambda\,(1-\varphi)\,(\tau_i-\tau_{i^*})-(1-\theta)\,\hat{\lambda}\,(1-\varphi)\,[1-\hat{\tau}_i(\theta)]}{[1-\hat{\tau}_i(\theta)]^2}\right)\right].$$
$$=\tau_{i^*}\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\left(\frac{1}{\left[1+\frac{1-\mathcal{C}(\lambda)}{\mathcal{C}'(\lambda)}\right]}\right)\left(\frac{\theta\left[\frac{\Lambda_{i^*}}{\sum_k\Lambda_k}\,(\tau_i-\tau_{i^*})\right]-(1-\hat{\tau}_i(\theta))}{(\tau_i-\tau_{i^*})\,(1-\hat{\tau}_i(\theta))}\right)$$
$$=\left(\frac{-\tau_{i^*}}{\tau_i-\tau_{i^*}}\right)\left(\frac{1-\gamma}{1-\phi-\gamma}\right)\left(\frac{1}{\left[1+\frac{1-\mathcal{C}(\hat{\lambda})}{\hat{\lambda}\mathcal{C}'(\hat{\lambda})}\right]}\right)\left(\frac{1-\tau_i-\theta\left[\frac{\Lambda_{i^*}}{\sum_k\Lambda_k}\,(\tau_i-\tau_{i^*})+\sum_k\tau_k\frac{\Lambda_k}{\sum_k\Lambda_k}-\tau_i\right]}{(1-\hat{\tau}_i(\theta))}\right).$$

Compare it to the elasticity of z^{PS}

$$\varepsilon_{\tau_{i^*}}^{z^{PS}}=\left(\frac{1-\gamma}{1-\phi+\gamma}\right)\left(\frac{-\tau_{i^*}}{\tau_i-\tau_{i^*}}\right)\frac{1}{\left[1+\frac{1-\mathcal{C}(\lambda)}{\mathcal{C}'(\lambda)}\right]},$$

and note that

$$\lim_{\theta\to 0}\varepsilon_{\tau_{i^*}}^{\hat{z}^{PS}}=\varepsilon_{\tau_{i^*}}^{z^{PS}}.$$

To show this, we have

$$\left(\frac{1-\tau_i-\theta\left[\frac{\Lambda_{i*}}{\sum_k \Lambda_k}\left(\tau_i-\tau_{i^*}\right)+\sum_k \tau_k \frac{\Lambda_k}{\sum_k \Lambda_k}-\tau_i\right]}{\left(1-\hat{\tau}_i(\theta)\right)}\right)=\left(\frac{\left(1-\hat{\tau}_i(\theta)\right)+\theta\left[\frac{\Lambda_{i*}}{\sum_k \Lambda_k}\left(\tau_i-\tau_{i^*}\right)\right]}{\left(1-\hat{\tau}_i(\theta)\right)}\right),$$

and under the assumption that sales to tax haven are negligible we have

$$\lim_{\Lambda_{i^*}\to 0}\left(\frac{\left(1-\hat{\tau}_i(\theta)\right)+\theta\left[\frac{\Lambda_{i*}}{\sum_k \Lambda_k}\left(\tau_i-\tau_{i^*}\right)\right]}{\left(1-\hat{\tau}_i(\theta)\right)}\right)=1.$$

Finally, we want to prove $\left|\varepsilon_{\tau_{i^*}}^{\hat{z}^{PS}}\right|<\left|\varepsilon_{\tau_{i^*}}^{z^{PS}}\right|$. It suffices to show that

$$\begin{aligned}
\frac{1}{\left(1+\frac{1-\mathcal{C}(\hat{\lambda})}{\hat{\lambda}\mathcal{C}'(\hat{\lambda})}\right)} &< \frac{1}{\left(1+\frac{1-\mathcal{C}(\lambda)}{\lambda\mathcal{C}'(\lambda)}\right)}\\
\left(1+\frac{1-\mathcal{C}\left(\hat{\lambda}\right)}{\hat{\lambda}\mathcal{C}'\left(\hat{\lambda}\right)}\right) &> \left(1+\frac{1-\mathcal{C}\left(\lambda\right)}{\lambda\mathcal{C}'\left(\lambda\right)}\right)\\
\frac{1-\mathcal{C}\left(\hat{\lambda}\right)}{\hat{\lambda}\mathcal{C}'\left(\hat{\lambda}\right)} &> \frac{1-\mathcal{C}\left(\lambda\right)}{\lambda\mathcal{C}'\left(\lambda\right)}.
\end{aligned}$$

We have

$$\begin{aligned}
\hat{\lambda} &< \lambda\\
\frac{1}{\hat{\lambda}} &> \frac{1}{\lambda},
\end{aligned}$$

but also

$$\begin{aligned}
\hat{\lambda} &< \lambda\\
\mathcal{C}\left(\hat{\lambda}\right) &< \mathcal{C}\left(\lambda\right)\\
1-\mathcal{C}\left(\hat{\lambda}\right) &> 1-\mathcal{C}\left(\lambda\right),
\end{aligned}$$

and also

$$\begin{aligned}
\mathcal{C}'\left(\lambda\right) &= -\log\left(1-\lambda\right)\\
\mathcal{C}''\left(\lambda\right) &= -\frac{(-1)}{1-\lambda}=\frac{1}{1-\lambda}>0.
\end{aligned}$$

Hence the marginal cost function is increasing in λ, therefore

$$\mathcal{C}'\left(\hat{\lambda}\right) < \mathcal{C}'\left(\lambda\right)$$
$$\frac{1}{\mathcal{C}'\left(\hat{\lambda}\right)} > \frac{1}{\mathcal{C}'\left(\lambda\right)}.$$

Therefore, we have that

$$\left(\frac{1}{\hat{\lambda}}\right)\left(1-\mathcal{C}\left(\hat{\lambda}\right)\right)\left(\frac{1}{\mathcal{C}'\left(\hat{\lambda}\right)}\right) > \left(\frac{1}{\lambda}\right)\left(1-\mathcal{C}\left(\lambda\right)\right)\left(\frac{1}{\mathcal{C}'\left(\lambda\right)}\right),$$

implying

$$\frac{1}{\left(1+\frac{1-\mathcal{C}(\hat{\lambda})}{\hat{\lambda}\mathcal{C}'(\hat{\lambda})}\right)} < \frac{1}{\left(1+\frac{1-\mathcal{C}(\lambda)}{\lambda\mathcal{C}'(\lambda)}\right)},$$

and hence

$$\left|\varepsilon_{\tau_{i^*}}^{\hat{z}^{PS}}\right| < \left|\varepsilon_{\tau_{i^*}}^{z^{PS}}\right|.$$

■

A.2.1 Alternative Assumption

Here, we assume that MNEs internalize the effect of changing z on the licensing fee $\vartheta_k(z)$ and solve for optimal z under different scenarios (FT, TP, and PS). We then prove Lemma 2 under this assumption. As before, we start from the optimal z.

Lemma 6 *The allocations of intangible capital are as follows:*

$$\hat{z}^{FT} = \left(\frac{\sum_k\left(1-\tau_k\right)\Lambda_k}{\left(1-\tau_i\right)p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}, \tag{88}$$

$$\hat{z}^{TP} = \left(\frac{\sum_k\left(1-\hat{\tau}_k\left(\theta\right)\right)\Lambda_k - \frac{\phi}{1-\gamma}\sum_k\left(1-\theta\right)\left(\tau_i-\tau_k\right)\Lambda_k}{\left(1-\hat{\tau}_i\left(\theta\right)\right)p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}, \tag{89}$$

$$\hat{z}^{PS} = \left(\frac{-\frac{\phi}{1-\gamma}\mathcal{C}\left(\hat{\lambda}\right)\sum_k\Lambda_k}{p_i} + \frac{\sum_k\left(1-\hat{\tau}_k\left(\theta\right)\right)\Lambda_k - \frac{\phi}{1-\gamma}\sum_k\left(1-\theta\right)\left(\tau_i-\tau_k\right)\Lambda_k + \hat{\lambda}\frac{\phi}{1-\gamma}\left(1-\theta\right)\left(\tau_i-\tau_{i^*}\right)\left(1-\varphi\right)\sum_k\Lambda_k}{\left(1-\hat{\tau}_i\left(\theta\right)\right)p_i}\right)^{\frac{1-\gamma}{1-\phi-\gamma}}. \tag{90}$$

Proof of Lemma 6. The proof follows the same procedure as the one of Lemma 4 ■

We are now ready to prove Lemma 2.

Proof of Lemma 2 under alternative assumption. We start from proving 1 from deriving a set of iff inequalities:

$$\frac{\sum_k \Lambda_k}{p_i} - \frac{\phi+\gamma-1}{1-\gamma}\frac{\sum_k (1-\theta)(\tau_i-\tau_k)\Lambda_k}{(1-\hat{\tau}_i(\theta))p_i} - \left(1+\frac{\phi+\gamma-1}{1-\gamma}\right)\frac{\sum_k \Lambda_k}{p_i}\left[\mathcal{C}\left(\hat{\lambda}\right) - \frac{\hat{\lambda}(1-\theta)(\tau_i-\tau_{i^*})(1-\varphi)}{(1-\hat{\tau}_i(\theta))}\right]$$
$$< \frac{\sum_k \Lambda_k}{p_i} - \frac{\phi+\gamma-1}{1-\gamma}\frac{\sum_k (\tau_i-\tau_k)\Lambda_k}{(1-\tau_i)p_i} - \left(1+\frac{\phi+\gamma-1}{1-\gamma}\right)\frac{\sum_k \Lambda_k}{p_i}\left[\mathcal{C}(\lambda) - \frac{\lambda(\tau_i-\tau_{i^*})(1-\varphi)}{(1-\tau_i)p_i}\right]$$
$$- \frac{1-\phi-\gamma}{1-\gamma}\frac{\sum_k (\tau_i-\tau_k)\Lambda_k}{(1-\tau_i)p_i} + \frac{\phi}{1-\gamma}\frac{\sum_k \Lambda_k}{p_i}\left[\mathcal{C}(\lambda) - \frac{\lambda(\tau_i-\tau_{i^*})(1-\varphi)}{(1-\tau_i)}\right]$$
$$< -\frac{1-\phi-\gamma}{1-\gamma}\frac{\sum_k (1-\theta)(\tau_i-\tau_k)\Lambda_k}{(1-\hat{\tau}_i(\theta))p_i} + \frac{\phi}{1-\gamma}\frac{\sum_k \Lambda_k}{p_i}\left[\mathcal{C}\left(\hat{\lambda}\right) - \frac{\hat{\lambda}(1-\theta)(\tau_i-\tau_{i^*})(1-\varphi)}{(1-\hat{\tau}_i(\theta))}\right].$$

We have proven before that $\mathcal{C}(\lambda) - \frac{\lambda(\tau_i-\tau_{i^*})(1-\varphi)}{(1-\tau_i)} < \mathcal{C}\left(\hat{\lambda}\right) - \frac{\hat{\lambda}(1-\theta)(\tau_i-\tau_{i^*})(1-\varphi)}{(1-\hat{\tau}_i(\theta))}$. It suffices to prove that

$$-\frac{1-\phi-\gamma}{1-\gamma}\frac{\sum_k (\tau_i-\tau_k)\Lambda_k}{(1-\tau_i)p_i} < -\frac{1-\phi-\gamma}{1-\gamma}\frac{\sum_k (1-\theta)(\tau_i-\tau_k)\Lambda_k}{(1-\hat{\tau}_i(\theta))p_i},$$

which simplifies to $1 > \sum_k \tau_k \frac{\Lambda_k}{\sum_i \Lambda_i}$. Thus, we have proven 1. We now prove 2:

$$\frac{d\hat{z}^{PS}}{d\theta} = \frac{1-\gamma}{1-\phi-\gamma}\left(\hat{z}^{PS}\right)^{\frac{\phi}{1-\gamma}}\frac{1}{p_i}\Bigg[-\frac{\phi}{1-\gamma}\mathcal{C}'\left(\hat{\lambda}\right)\frac{d\hat{\lambda}}{d\theta}\sum_k \Lambda_k + \left(\frac{1-\theta}{(1-\hat{\tau}_i(\theta))^2}\hat{\tau}_i'(\theta) - \frac{1}{1-\hat{\tau}_i(\theta)}\right)\cdot$$
$$\left(\frac{\phi}{1-\gamma}\hat{\lambda}(\tau_i-\tau_{i^*})(1-\varphi) - \frac{\phi+\gamma-1}{1-\gamma}\sum_k(\tau_i-\tau_k)\Lambda_k\right) +$$
$$\frac{d\hat{\lambda}}{d\theta}\frac{\phi}{1-\gamma}\frac{(1-\theta)(\tau_i-\tau_{i^*})(1-\varphi)}{1-\hat{\tau}_i(\theta)}\sum_k \Lambda_k\Bigg]$$
$$= \frac{1-\gamma}{1-\phi-\gamma}\left(\hat{z}^{PS}\right)^{\frac{\phi}{1-\gamma}}\frac{1}{p_i}\left(\frac{\sum_k \tau_k\frac{\Lambda_k}{\sum_i \Lambda_i} - 1}{(1-\hat{\tau}_i(\theta))^2}\right)\left(\frac{\phi}{1-\gamma}\lambda(\tau_i-\tau_{i^*})(1-\varphi) + \frac{1-\phi-\gamma}{1-\gamma}\sum_k(\tau_i-\tau_k)\Lambda_k\right).$$

We have shown that

$$\sum_k \tau_k \frac{\Lambda_k}{\sum_i \Lambda_i} - 1 < 0.$$

The other terms are all positive, thus we have proven 2. Now to prove 3 we can show that

$$\frac{\partial\hat{z}^{PS}}{\partial\tau_{i^*}} = \left(\frac{1-\gamma}{1-\phi-\gamma}\right)\left(\hat{z}^{PS}\right)^{\frac{\phi}{1-\gamma}}\frac{1}{p_i}\Bigg\{-\frac{\phi}{1-\gamma}\mathcal{C}'\left(\hat{\lambda}\right)\frac{\partial\hat{\lambda}}{\partial\tau_{i^*}}\sum_k \Lambda_k + \frac{\phi}{1-\gamma}\frac{\sum_k \Lambda_k}{(1-\hat{\tau}_i(\theta))^2}\cdot$$
$$\left[\left(\frac{\partial\hat{\lambda}}{\partial\tau_{i^*}}(1-\varphi)(\tau_i-\tau_{i^*})(1-\theta)\sum_k \Lambda_k - (1-\theta)\hat{\lambda}(1-\varphi)\right)(1-\hat{\tau}_i(\theta)) + \left(\theta\frac{\Lambda_{i^*}}{\sum_k \Lambda_k}\right)(1-\theta)\hat{\lambda}(1-\varphi)(\tau_i-\tau_{i^*})\right]$$
$$-\frac{1-\phi-\gamma}{1-\gamma}(1-\theta)\Lambda_{i^*}\frac{1-\tau_i}{(1-\hat{\tau}_i(\theta))^2}\Bigg\}.$$

We have shown in previous proof that the sum of first two terms in the big bracket is negative. It's clear that the last term is also negative. Hence we have proven 3, that is $\frac{\partial\hat{z}^{PS}}{\partial\tau_{i^*}} < 0$. ■

B Quantitative model

B.1 Firm's problem with no transfer pricing or profit shifting

B.1.1 Scale choice: the parent division

We start from the parent division of a firm $\omega \in \Omega_i$'s scale choice here. A parent division that produces for the domestic market and exports to a set of J_X regions chooses its scale and how to allocate its output across its markets. Note that this problem nests the problem for firms only producing for the domestic markets when $J_X = \emptyset$. The parent division's problem can then be written as

$$\pi_i^D(a, z; J_X) = \max_{q_{ii}, (q_{ij})_{j \in J_X}, \ell} \left\{ p_{ii}(q_{ii}) q_{ii} + \sum_{j \in J_X} p_{ij}(q_{ij}^X) q_{ij}^X - W_i \ell \right\},$$
$$\text{s.t} \quad q_{ii} + \sum_{j \in J_X} \xi_{ij} q_{ij}^X = y_i = A_i a (N_i z)^\gamma \ell^\phi.$$

The first-order conditions are

$$[q_{ij}] \quad \frac{\varrho - 1}{\varrho} P_j Q_j^{\frac{1}{\varrho}} q_{ij}^{-\frac{1}{\varrho}} = \lambda \xi_{ij},$$
$$[\ell] \quad W_i = \lambda \phi A_i a (N_i z)^\gamma \ell^{\phi - 1},$$

where $\xi_{ii} = 1$. Rearrange to get

$$\frac{\varrho - 1}{\varrho} P_j Q_j^{\frac{1}{\varrho}} q_{ij}^{-\frac{1}{\varrho}} = \frac{\tau_{ij} W_i}{\phi A_i a (N_i z)^\gamma \ell^{\phi - 1}}.$$

Then

$$q_{ij} = \left[\frac{\phi(\theta - 1)}{\theta} \right]^\theta \left[\frac{P_j Q_j^{\frac{1}{\theta}} A_i a (N_i z)^\gamma \ell^{\phi - 1}}{\tau_{ij} W_i} \right]^\theta = \left[\frac{P_j Q_j^{\frac{1}{\theta}}}{\tau_{ij}} \right]^\theta \left[\frac{\phi(\theta - 1)}{\theta} \right]^\theta \left[\frac{A_i a (N_i z)^\gamma \ell^{\phi - 1}}{W_i} \right]^\theta.$$

Plugging this back into the resource constraint, we have

$$\left[P_i^\theta Q_i + \sum_{j \in J_X} P_j^\theta \tau_j^{1-\theta} Q_j \right] \left[\frac{\phi(\theta - 1)}{\theta} \right]^\theta \left[\frac{A_i a (N_i z)^\gamma \ell^{\phi - 1}}{W_i} \right]^\theta = A_i a (N_i z)^\gamma \ell^\phi,$$

which simplifies to

$$\left[P_i^\theta Q_i + \sum_{j \in J_X} P_j^\theta \tau_j^{1-\theta} Q_j \right] \left[\frac{\phi(\theta - 1)}{\theta} \right]^\theta W_i^{-\theta} (A_i a)^{\theta - 1} (N_i z)^{\gamma(\theta - 1)} = \ell^{\phi + \theta - \theta\phi}.$$

We can solve the optimal labor choice

$$\ell = \left\{ \left[P_i^\varrho Q_i + \sum_{j \in J_X} P_j^\varrho \tau_j^{1-\varrho} Q_j \right] \left[\frac{\phi(\varrho - 1)}{\varrho} \right]^\varrho W_i^{-\varrho} (A_i a)^{\varrho - 1} (N_i z)^{\gamma(\varrho - 1)} \right\}^{\frac{1}{\phi + \varrho - \varrho\phi}}. \tag{91}$$

We can use the equations above to compute q_{ij}, p_{ij}, and $\pi_i^D(a,z;J_X)$.

B.1.2 Scale choice: foreign subsidiaries

Foreign subsidiaries are similar to domestic-only firms. They just choose scale to maximize profits from selling to the host market given the demand curve and production technology. The only difference is the presence of the FDI barrier σ_{ij}. The foreign subsidiary's problem is

$$\begin{aligned}\pi_{ij}^F(a,z) &= \max_{q,\ell}\ p_{ij}(q)q - W_i\ell \\ &= \max_{\ell}\ P_jQ_j^{\frac{1}{\varrho}}\left(\sigma_{ij}A_ja\right)^{\frac{\varrho-1}{\varrho}}\left(N_jz\right)^{\gamma\frac{\varrho-1}{\varrho}}\ell^{\phi\frac{\varrho-1}{\varrho}} - W_j\ell.\end{aligned}$$

The FOC is

$$\phi\frac{\varrho-1}{\varrho}P_jQ_j^{\frac{1}{\varrho}}\left(\sigma_{ij}A_ja\right)^{\frac{\varrho-1}{\varrho}}\left(N_jz\right)^{\gamma\frac{\varrho-1}{\varrho}}\ell^{\phi\frac{\varrho-1}{\varrho}-1} = W_j.$$

The optimal ℓ is then

$$\ell = \left\{\left[\frac{\phi(\varrho-1)}{\varrho}\right]^{\varrho}(P_j/W_j)^{\varrho}Q_j\left(\sigma_{ij}A_ja\right)^{\varrho-1}\left(N_jz\right)^{\gamma(\varrho-1)}\right\}^{\frac{1}{\phi+\varrho-\phi\varrho}}. \tag{92}$$

We can use this to compute $q_{ij} = y_j = \sigma_{ij}A_ja(N_jz)^{\gamma}\ell^{\phi}$, $p_{ij} = P_jQ_j^{\frac{1}{\varrho}}q_{ij}^{-\frac{1}{\varrho}}$, and $\pi_j^F(a,z)$.

B.1.3 Technology choice

Now that we have $\pi_i^D(a,z;J_X)$ of the parent division and $\pi_{ij}^F(a,z)$ of foreign affiliates, $j \in J_F$, we can determine how much R&D to do taking J_F and J_X as given. Note that we can ignore the fixed costs of exporting and FDI for now:

$$\hat{d}_i(a;J_X,J_F) = \max_z\left\{(1-\tau_i)\left[\pi_i^D(a,z;J_X) - W_iz/A_i\right] + \sum_{j\in J_F}(1-\tau_j)\pi_{ij}^F(a,z)\right\}.$$

Using the solutions for labor, the parent corporation's output can be written as

$$\begin{aligned}y_{ii} &= Aa(N_iz)^{\gamma}\left\{\left[P_i^{\varrho}Q_i + \sum_{j\in J_X}P_j^{\varrho}\tau_j^{1-\varrho}Q_j\right]\left[\frac{\phi(\varrho-1)}{\varrho}\right]^{\varrho}W^{-\varrho}(A_ia)^{\varrho-1}(N_iz)^{\gamma(\varrho-1)}\right\}^{\frac{\phi}{\phi+\varrho-\varrho\phi}} \\ &= \left\{\left[P_i^{\varrho}Q_i + \sum_{j\in J_X}P_j^{\varrho}\tau_j^{1-\varrho}Q_j\right]\left[\frac{\phi(\varrho-1)}{\varrho}\right]^{\varrho}W^{-\varrho}\right\}^{\frac{\phi}{\phi+\varrho-\varrho\phi}}(A_ia)^{\frac{\varrho}{\phi+\varrho-\phi\varrho}}(N_iz)^{\frac{\gamma\varrho}{\phi+\varrho-\phi\varrho}}.\end{aligned}$$

We can use the FOC for q_{ij} to write

$$\frac{P_jQ_j^{\frac{1}{\varrho}}q_{ij}^{-\frac{1}{\varrho}}}{P_kQ_k^{\frac{1}{\varrho}}q_{ik}^{-\frac{1}{\varrho}}} = \frac{\tau_{ij}}{\tau_{ik}} \Rightarrow q_{ij} = \left(\frac{\tau_{ij}}{\tau_{ik}}\right)^{-\varrho}\left(\frac{P_j}{P_k}\right)^{\varrho}\left(\frac{Q_j}{Q_k}\right)q_{ik}.$$

Set $k = i$ and combine this with the resource constraint to get

$$q_{ii} + \sum_{j \in J_X} \tau_{ij}^{1-\varrho} \left(\frac{P_j}{P_i}\right)^{\varrho} \left(\frac{Q_j}{Q_i}\right) q_{ii} = y_{ii} \Rightarrow q_{ii} = \left(\frac{1}{1 + \sum_{j \in J_X} \tau_{ij}^{1-\varrho} \left(\frac{P_j}{P_i}\right)^{\varrho} \left(\frac{Q_j}{Q_i}\right)}\right) y_{ii} = \bar{Q}_{ii} y_{ii}.$$

We can then write

$$q_{ij} = \tau_{ij}^{-\varrho} \left(\frac{P_j}{P_i}\right)^{\varrho} \left(\frac{Q_j}{Q_i}\right) \bar{Q}_{ii} y_{ii} = \bar{Q}_{ij} y_{ii}.$$

Then domestic parent revenues are

$$\begin{aligned} p_{ii} q_{ii} + \sum_{j \in J_X} p_{ij} q_{ij} &= P_i Q_i^{\frac{1}{\varrho}} q_{ii}^{\frac{\varrho-1}{\varrho}} + \sum_{j \in J_X} P_j Q_j^{\frac{1}{\varrho}} q_{ij}^{\frac{\varrho-1}{\varrho}} \\ &= \left[P_i Q_i^{\frac{1}{\varrho}} \bar{Q}_{ii}^{\frac{\varrho-1}{\varrho}} + \sum_{j \in J_X} P_j Q_j^{\frac{1}{\varrho}} \bar{Q}_{ij}^{\frac{\varrho-1}{\varrho}} \right] \\ &\qquad \times \left\{ \left[P_i^{\varrho} Q_i + \sum_{j \in J_X} P_j^{\varrho} \tau_j^{1-\varrho} Q_j \right] \left[\frac{\phi(\varrho-1)}{\varrho} \right]^{\varrho} W^{-\varrho} \right\}^{\frac{\varrho-1}{\varrho} \frac{\phi}{\phi+\varrho-\varrho\phi}} \\ &\qquad \times (A_i a)^{\frac{\varrho-1}{\phi+\varrho-\phi\varrho}} N_i^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} \\ &= \bar{R}_{ii} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}. \end{aligned} \tag{93}$$

Domestic parent costs are

$$\begin{aligned} W_i \ell + W_i z / A_i &= W_i \left\{ \left[P_i^{\varrho} Q_i + \sum_{j \in J_X} P_j^{\varrho} \tau_j^{1-\varrho} Q_j \right] \left[\frac{\phi(\varrho-1)}{\varrho} \right]^{\varrho} W^{-\varrho} (A_i a)^{\varrho-1} (N_i z)^{\gamma(\varrho-1)} \right\}^{\frac{1}{\phi+\varrho-\varrho\phi}} + W_i z / A_i \\ &= \bar{C}_{ii} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\varrho\phi}} + W_i z / A_i. \end{aligned}$$

Foreign affiliate revenues are

$$\begin{aligned} p_{ij} q_{ij} &= P_j Q_j^{\frac{1}{\varrho}} q_{ij}^{\frac{\varrho-1}{\varrho}} \\ &= \left[P_j Q_j^{\frac{1}{\varrho}} \right] \left[(P_j / W_j) \frac{\phi(\varrho-1)}{\varrho} \right]^{\frac{\phi(\varrho-1)}{\phi+\varrho-\phi\varrho}} Q_j^{\frac{\varrho-1}{\varrho} \frac{\phi}{\phi+\varrho-\phi\varrho}} (A_j \sigma_{ij} a)^{\frac{\varrho-1}{\phi+\varrho-\phi\varrho}} N_j^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} \\ &= \bar{R}_{ij} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}. \end{aligned} \tag{94}$$

Foreign affiliate costs are

$$\begin{aligned} W_j \ell &= W_j \left\{ \left[\frac{\phi(\varrho-1)}{\varrho} \right]^{\varrho} (P_j / W_j)^{\varrho} Q_j \, (A_j \sigma_{ij} a)^{\varrho-1} \, (N_j z)^{\gamma(\varrho-1)} \right\}^{\frac{1}{\phi+\varrho-\phi\varrho}} \\ &= \bar{C}_{ij} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}. \end{aligned}$$

Total net revenues are

$$(1-\tau_i)p_{ii}q_{ii} + \sum_{j\in J_X}(1-\tau_j)p_{ij}q_{ij} + \sum_{j\in J_F} p_{ij}q_{ij} = (1-\tau_i)\bar{R}_{ii}z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} + \sum_{j\in J_F}(1-\tau_j)\bar{R}_{ij}z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}.$$

Total costs are

$$(1-\tau_i)(W_i\ell_{ii} + W_i z/A_i) + \sum_{j\in J_F}(1-\tau_j)W_j\ell_{ij} = \left[(1-\tau_i)\bar{C}_{ii} + \sum_{j\in J_F}(1-\tau_j)C_{ij}\right] z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\varrho\phi}} + (1-\tau_i)W_i z/A_i.$$

We can write the objective function as

$$\left[(1-\tau_i)(\bar{R}_{ii} - \bar{C}_{ii}) + \sum_{j\in J_F}(1-\tau_j)\left(\bar{R}_{ij} - \bar{C}_{ij}\right)\right] z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} - (1-\tau_i)W_i z/A_i.$$

The FOC is

$$\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)\left[(1-\tau_i)(\bar{R}_{ii} - \bar{C}_{ii}) + \sum_{j\in J_F}(1-\tau_j)\left(\bar{R}_{ij} - \bar{C}_{ij}\right)\right] z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}-1} = (1-\tau_i)W_i/A_i.$$

Then the optimal choice of z is

$$z = \left\{\left(\frac{\phi+\varrho-\phi\varrho}{\gamma(\varrho-1)}\right)\left[\frac{(1-\tau_i)W_i/A_i}{(1-\tau_i)\left(\bar{R}_{ii} - \bar{C}_{ii}\right) + \sum_{j\in J_F}(1-\tau_j)\left(\bar{R}_{ij} - \bar{C}_{ij}\right)}\right]\right\}^{\frac{\phi+\varrho-\phi\varrho}{\gamma\varrho+\phi\varrho-\gamma-\phi-\varrho}}.$$

B.1.4 Market choice

Now that we have $\hat{d}_i(a; J_X, J_F), \forall J_X, J_F$, we can decide where to export and where to operate foreign subsidiaries:

$$d_i(a) = \max_{J_X, J_F}\left\{\hat{d}_i(a; J_X, J_F) - W_i\left(\sum_{j\in J_X}\kappa_{jX} - \sum_{j\in J_F}\kappa_{jF})\right)\right\}. \tag{95}$$

This is a combinatorial discrete choice problem discussed in Arkolakis et al. (2021), as a firm's exporting and FDI choices interdependent. This problem is hard to solve since the number of potential decision sets grows exponentially in the number of regions. We limit the number of regions in the quantitative model to ease the computational burden.

B.2 Firm's problem with transfer pricing

Here, we solve the optimal non-rival technology allocation z in the environment with transfer pricing, taking J_X and J_F as given. We define total profit earned by a firm in this scenario as *transfer-pricing profit*, d_i^{TP}:

$$d_i^{TP}(a) = \max_{J_X, J_F}\left\{\hat{d}_i^{TP}(a; J_X, J_F) - W_i\left(\sum_{j\in J_X}\kappa_{jX} - \sum_{j\in J_F}\kappa_{jF})\right)\right\}, \tag{96}$$

where

$$\hat{d}_i^{TP}(a;J_X,J_F)=\max_z\Bigg\{(1-\tau_i)\left[\pi_i^D(a,z;J_X)+(\sum_{j\in J_F}\vartheta_j(z)-W_i/A_i)z\right]$$
$$+\sum_{j\in J_F}(1-\tau_j)(\pi_{ij}^F(a,z)-\vartheta_j(z)\cdot z)\Bigg\}. \tag{97}$$

Taking J_X and J_F as given, each firm chooses z to maximize $\hat{d}_i^{TP}(a;J_X,J_F)$. We can write the objective function as

$$\max_z\Bigg\{(1-\tau_i)\left[(\bar{R}_{ii}-\bar{C}_{ii})+\sum_{j\in J_F}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]$$
$$+\sum_{j\in J_F}(1-\tau_j)\left(1-\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\Bigg\}z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}-(1-\tau_i)W_iz/A_i.$$

The FOC is

$$\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)\Bigg\{(1-\tau_i)\left[\bar{R}_{ii}-\bar{C}_{ii}+\sum_{j\in J_F}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]$$
$$+\sum_{j\in J_F}(1-\tau_j)\left(1-\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\Bigg\}z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}-1}=(1-\tau_i)W_i/A_i.$$

Then the optimal choice of z is

$$z=\left\{\left(\frac{\phi+\varrho-\phi\varrho}{\gamma(\varrho-1)}\right)\left[\frac{(1-\tau_i)W_i/A_i}{DENOM^{TP}}\right]\right\}^{\frac{\phi+\varrho-\phi\varrho}{\gamma\varrho+\phi\varrho-\gamma-\phi-\varrho}},$$

where

$$DENOM^{TP}=(1-\tau_i)\left[(\bar{R}_{ii}-\bar{C}_{ii})+\sum_{j\in J_F}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]$$
$$+\sum_{j\in J_F}(1-\tau_j)\left(1-\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}).$$

B.3 Firm's problem with transfer pricing and profit shifting

Here, we solve the optimal non-rival technology allocation z and profit shifting shares λ_{TH} and λ_{LT} in the environment with transfer pricing and profit shifting, taking J_X and J_F as given. This problem nests the one with only transfer pricing and no profit shifting simply by setting λ and $C(\lambda)$ to zero for both LT and TH. Here we solve for the full problem where λ_{LT} and $\lambda_{TH}>0$. It is easier to rewrite $\hat{d}^{PS}(a)$ as:

$$d_i^{PS}(a)=\max_{J_X,J_F}\left\{\hat{d}_i^{PS}(a;J_X,J_F)-W_i\left(\sum_{j\in J_X}\kappa_{jX}-\sum_{j\in J_F}\kappa_{jF}-\kappa_{iTH}\mathbb{1}(\lambda_{TH}>0)\right)\right\},$$

where

$$\hat{d}_i^{PS} = \max_{z,\lambda_{LT},\lambda_{TH}} \Bigg\{ (1-\tau_i)\Big[\pi_i^D(a,z;J_X) + \Big(\varphi_{iLT}\lambda_{LT}\nu_i(z) + \varphi_{iTH}\lambda_{TH}\nu_i(z) - (\lambda_{LT}+\lambda_{TH})\vartheta_i(z) + (1-\lambda_{LT}-\lambda_{TH})\sum_{j\in J_F}\vartheta_j(z) - W_i/A_i - C_{i,LT}(\lambda_{LT})\nu_i(z) - C_{i,TH}(\lambda_{TH})\nu_i(z)\Big)z\Big]$$
$$+(1-\tau_{LT})\Big[\pi_{i,LT}^F(a,z) + \Big(\lambda_{LT}\sum_{j\in J_F\cup\{i\}\setminus\{LT\}}\vartheta_j(z) - (1-\lambda_{LT})\vartheta_{iLT}(z) - \varphi_{iLT}\lambda_{LT}\nu_i(z)\Big)z\Big]$$
$$+(1-\tau_{TH})\Big[\Big(\lambda_{TH}\sum_{j\in J_F\cup\{i\}}\vartheta_j(z) - \varphi_{iTH}\lambda_{TH}\nu_i(z)\Big)z\Big]$$
$$+ \sum_{j\in J_F\setminus\{LT\}}(1-\tau_j)\left[\pi_{ij}^F(a,z) - \vartheta_j(z)z\right] \Bigg\}.$$

Substituting in the optimal scale choices specified in equation 91 and 92 and letting $\lambda = \lambda_{TH} + \lambda_{LT}$, we can write $\hat{d}_i^{PS}$ as

$$\max_{z,\lambda_{TH},\lambda_{LT}} \Bigg\{ (1-\tau_i)\left[\left(1-\lambda\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)\right)(\bar{R}_{ii}-\bar{C}_{ii}) + (1-\lambda)\sum_{j\in J_F}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]$$
$$-(1-\tau_i)\left[(C_{i,LT}(\lambda_{LT}) + C_{i,TH}(\lambda_{TH}) - \varphi_{iLT}\lambda_{LT} - \varphi_{iTH}\lambda_{TH})\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]$$
$$+(1-\tau_{LT})\Bigg[\left(1-(1-\lambda_{LT})\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)\right)(\bar{R}_{i,LT}-\bar{C}_{i,LT}) + \lambda_{LT}\sum_{j\in J_F\cup\{i\}\setminus\{LT\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\Bigg]$$
$$-(1-\tau_{LT})\left[\varphi_{iLT}\lambda_{LT}\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]$$
$$+(1-\tau_{TH})\left[\lambda_{TH}\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}) - \varphi_{iTH}\lambda_{TH}\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]$$
$$+\sum_{j\in J_F\setminus\{LT\}}(1-\tau_j)\left[\left(1-\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\right]\Bigg\} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} - (1-\tau_i)W_i z/A_i.$$

And further simplifying

$$\max_{z,\lambda_{TH},\lambda_{LT}} \Bigg\{ \sum_{j\in J_F\cup\{i\}} (1-\tau_j)(\bar{R}_{ij}-\bar{C}_{ij}) - \sum_{j\in J_F\cup\{i\}} (\tau_i-\tau_j)\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+ (\tau_i-\tau_{LT})\lambda_{LT} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+ (\tau_i-\tau_{TH})\lambda_{TH} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+ (\tau_{LT}-\tau_i)\varphi_{iLT}\lambda_{LT} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+ (\tau_{TH}-\tau_i)\varphi_{iTH}\lambda_{TH} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$- (1-\tau_i)\left(C_{i,LT}(\lambda_{LT})+C_{i,TH}(\lambda_{TH})\right) \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\Bigg\} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} - (1-\tau_i)W_i z/A_i.$$

Recall that the λ values can be solved independent of z:

$$\lambda_{LT} = \left(\mathcal{C}'_{i,LT}\right)^{-1}\left[(1-\varphi_{iLT})\frac{(\tau_i-\tau_{LT})}{1-\tau_i}\right],$$
$$\lambda_{TH} = \left(\mathcal{C}'_{i,TH}\right)^{-1}\left[(1-\varphi_{iTH})\frac{(\tau_i-\tau_{LT})}{1-\tau_i}\right].$$

The FOC for z is

$$(1-\tau_i)W_i/A_i = \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}-1} \Bigg\{ \sum_{j\in J_F\cup\{i\}} (1-\tau_j)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$- \sum_{j\in J_F\cup\{i\}} (\tau_i-\tau_j)\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+(\tau_i-\tau_{LT})\lambda_{LT} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+(\tau_{LT}-\tau_i)\varphi_{iLT}\lambda_{LT} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+(\tau_i-\tau_{TH})\lambda_{TH} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$+(\tau_{TH}-\tau_i)\varphi_{iTH}\lambda_{TH} \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})$$
$$-(1-\tau_i)\left(C_{i,LT}(\lambda_{LT})+C_{i,TH}(\lambda_{TH})\right) \sum_{j\in J_F\cup\{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\Bigg\}.$$

We can solve the optimal z as:

$$z = \left\{ \left(\frac{\phi + \varrho - \phi\varrho}{\gamma(\varrho - 1)} \right) \left[\frac{(1-\tau_i)\, W_i/A_i}{DENOM^{PS}} \right] \right\}^{\frac{\phi+\varrho-\phi\varrho}{\gamma\varrho+\phi\varrho-\gamma-\phi-\varrho}},$$

where $DENOM^{PS}$ is the stuff inside the big brackets above.

B.4 Profit allocation rule

As before, we solve for the full problem where $\lambda_{LT} > 0$ and $\lambda_{TH} > 0$. It's easier to state the firm's problem as:

$$d_i^{PR}(a) = \max_{z, J_X, J_F, \lambda_{TL}, \lambda_{TH}} \left\{ \hat{d}_i^{PR}(a; J_X, J_F) - W_i \left(\sum_{j \in J_X} \kappa_{jX} - \sum_{j \in J_F} \kappa_{jF} - \kappa_{iTH} \mathbb{1}(\lambda_{TH} > 0) \right) \right\}.$$

Each firm, taking J_X and J_F as given, chooses z and λ to maximize

$$\hat{d}_i^{PR}(a; J_X, J_F, \varrho) = \max_{z, \lambda_{TH}, \lambda_{LT}} \left\{ \sum_{j \in \{i\} \cup J_X \cup J_F} \left(\pi_j^{PR}(a, z) - \tau_j T_j \right) \right\}, \tag{98}$$

where

$$\begin{aligned}
\hat{d}_i^{PR} = \max_{z, \lambda_{LT}, \lambda_{TH}} \Bigg\{ \Bigg[\pi_i^D(a, z; J_X) + \Bigg(& \varphi_{iLT} \lambda_L T \nu_i(z) + \varphi_{iTH} \lambda_{TH} \nu_i(z) \\
& - (\lambda_{LT} + \lambda_{TH}) \vartheta_i(z) + (1 - \lambda_{LT} - \lambda_{TH}) \sum_{j \in J_F} \vartheta_j(z) \\
& - W_i/A_i - C_{i,LT}(\lambda_{LT}) \nu_i(z) - C_{i,TH}(\lambda_{TH}) \nu_i(z) \Bigg) z - \tau_i T_i \Bigg] \\
+ \Bigg[\pi_{i,LT}^F(a, z) + \Bigg(\lambda_{LT} & \sum_{j \in J_F \cup \{i\} \setminus \{LT\}} \vartheta_j(z) - (1 - \lambda_{LT}) \vartheta_{iLT}(z) - \varphi_{iLT} \lambda_{LT} \nu_i(z) \Bigg) z - \tau_{LT} T_{LT} \Bigg] \\
+ \Bigg[\Bigg(\lambda_{TH} \sum_{j \in J_F \cup \{i\}} & \vartheta_j(z) - \varphi_{iTH} \lambda_{TH} \nu_i(z) \Bigg) z - \tau_{TH} T_{TH} \Bigg] \\
+ \sum_{j \in J_F \setminus \{LT\}} & \left[\pi_{ij}^F(a, z) - \vartheta_j(z) z - \tau_j T_j \right] + \sum_{j \in J_X \setminus J_F} \left[-\tau_j T_j \right] \Bigg\}.
\end{aligned}$$

T_j is the tax base in region j

$$\begin{aligned}
T_j =& \Pi_j^r + (1-\theta) \cdot \Pi_j^R + \theta \cdot \frac{R_j}{\sum_k R_k} \cdot \Pi^R \\
&= \vartheta R_j + (1-\theta) \cdot (\pi_j(a, z; J_X) - \mu R_j) + \theta \cdot \frac{R_j}{\sum_k R_k} \cdot \sum_{k \in \{i\} \cup J_X \cup J_F} (\pi_k(a, z; J_X) - \mu R_k) \\
&= (1-\theta) \cdot \pi_j(a, z; J_X) + \theta \cdot \frac{R_j}{\sum_k R_k} \cdot \sum_{k \in \{i\} \cup J_X \cup J_F} \pi_k(a, z; J_X).
\end{aligned}$$

Profit π_j is the profit earned in region j and it is zero if the firm does not operate in the region. Revenue earned in region j, denoted as R_j, include sales of both goods produced locally (by parent division or FDI) and goods exported to the region. Formally:

$$
\begin{aligned}
R_i &= p_{ii}\left(q_{ii}\right) q_{ii}, \\
R_j &= p_{ij}^F\left(q_{ij}\right) q_{ij}, j \in J_F, j \notin J_X, \\
R_j &= p_{ij}^X\left(q_{ij}^X\right) q_{ij}^X, j \in J_X, j \notin J_F, \\
R_j &= p_{ij}^F\left(q_{ij}\right) q_{ij} + p_{ij}^X\left(q_{ij}^X\right) q_{ij}^X, j \in J_X \cap J_F, \\
R_j &= 0, \quad j \notin \{i\} \cup J_F \cup J_X.
\end{aligned}
$$

Thus, we can rewrite firm's problem as:

$$
\hat{d}_i^{PR} = \max_{z, \lambda_{TH}, \lambda_{LT}} \left\{ \sum_{j \in \{i\} \cup J_X \cup J_F} \left(\left(1 - \tau_j(1-\theta)\right) \pi_j(a, z; J_X) - \tau_j \theta \cdot \frac{R_j}{\sum_j R_j} \cdot \sum_k \pi_k(a, z; J_X) \right) \right\}.
$$

Further, substituting in π_i and denoting $\lambda = \lambda_{TH} + \lambda_{LT}$, we get

$$
\begin{aligned}
\max_{z, \lambda_{TH}, \lambda_{LT}} \Bigg\{ & (1-(1-\theta)\tau_i) \left[\left(1 - \lambda \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)\right) (\bar{R}_{ii} - \bar{C}_{ii}) + (1-\lambda) \sum_{j \in J_F} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) (\bar{R}_{ij} - \bar{C}_{ij}) \right] \\
& -(1-(1-\theta)\tau_i) \left[\left(C_{i,LT}(\lambda_{LT}) + C_{i,TH}(\lambda_{TH}) - \varphi_{iLT}\lambda_{LT} - \varphi_{iTH}\lambda_{TH}\right) \sum_{j \in J_F \cup \{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) (\bar{R}_{ij} - \bar{C}_{ij}) \right] \\
& +(1-(1-\theta)\tau_{LT}) \Bigg[\left(1 - (1-\lambda_{LT}) \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)\right) (\bar{R}_{i,LT} - \bar{C}_{i,LT}) + \\
& \qquad \lambda_{LT} \sum_{j \in J_F \cup \{i\} \backslash \{LT\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) (\bar{R}_{ij} - \bar{C}_{ij}) \Bigg] \\
& -(1-(1-\theta)\tau_{LT}) \left[\varphi_{iLT}\lambda_{LT} \sum_{j \in J_F \cup \{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) (\bar{R}_{ij} - \bar{C}_{ij}) \right] \\
& +(1-(1-\theta)\tau_{TH}) \left[\lambda_{TH} \sum_j \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) (\bar{R}_{ij} - \bar{C}_{ij}) - \varphi_{iTH}\lambda_{TH} \sum_{j \in J_F \cup \{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) (\bar{R}_{ij} - \bar{C}_{ij}) \right] \\
& + \sum_{j \in J_F \backslash \{LT\}} (1-\tau_j) \left[\left(1 - \frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) (\bar{R}_{ij} - \bar{C}_{ij}) \right] \Bigg\} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} - (1-(1-\theta)\tau_i) W_i z / A_i \\
& - \sum_{j \in \{i\} \cup J_X \cup J_F} \tau_j \theta \cdot \frac{R_j}{\sum_j R_j} \cdot \sum_k \pi_k(a, z; J_X).
\end{aligned}
$$

Here we define $\tilde{R}_{ij}$ as the revenue shifter in region j for firms from region i, depending on region j is served. These terms are defined analogously of $\bar{R}_{ij}$ in equations 93 and 94.

$$\tilde{R}_{ii} = P_i Q_i^{\frac{1}{\varrho}} \bar{Q}_{ii}^{\frac{\varrho-1}{\varrho}} \left\{ \left[P_i^{\varrho} Q_i + \sum_{j \in J_X} P_j^{\varrho} \tau_j^{1-\varrho} Q_j \right] \left[\frac{\phi(\varrho-1)}{\varrho} \right]^{\varrho} W^{-\varrho} \right\}^{\frac{\varrho-1}{\varrho} \frac{\phi}{\phi+\varrho-\varrho\phi}} (A_i)^{\frac{\varrho-1}{\phi+\varrho-\phi\varrho}} N_i^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}},$$

$$\tilde{R}_{ij} = P_j Q_j^{\frac{1}{\varrho}} \bar{Q}_{ij}^{\frac{\varrho-1}{\varrho}} \left\{ \left[P_i^{\varrho} Q_i + \sum_{j \in J_X} P_j^{\varrho} \tau_j^{1-\varrho} Q_j \right] \left[\frac{\phi(\varrho-1)}{\varrho} \right]^{\varrho} W^{-\varrho} \right\}^{\frac{\varrho-1}{\varrho} \frac{\phi}{\phi+\varrho-\varrho\phi}} (A_i)^{\frac{\varrho-1}{\phi+\varrho-\phi\varrho}} N_i^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}, \quad j \in J_X, j \notin J_F,$$

$$\tilde{R}_{ij} = \left[P_j Q_j^{\frac{1}{\varrho}} \right] \left[(P_j/W_j) \frac{\phi(\varrho-1)}{\varrho} \right]^{\frac{\phi(\varrho-1)}{\phi+\varrho-\phi\varrho}} Q_j^{\frac{\varrho-1}{\varrho} \frac{\phi}{\phi+\varrho-\phi\varrho}} (A_j \sigma_{ij})^{\frac{\varrho-1}{\phi+\varrho-\phi\varrho}} N_j^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}, \quad j \in J_F, j \notin J_X,$$

$$\tilde{R}_{ij} = \left[P_j Q_j^{\frac{1}{\varrho}} \right] \left[(P_j/W_j) \frac{\phi(\varrho-1)}{\varrho} \right]^{\frac{\phi(\varrho-1)}{\phi+\varrho-\phi\varrho}} Q_j^{\frac{\varrho-1}{\varrho} \frac{\phi}{\phi+\varrho-\phi\varrho}} (A_j \sigma_{ij})^{\frac{\varrho-1}{\phi+\varrho-\phi\varrho}} N_j^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}$$
$$+ P_j Q_j^{\frac{1}{\varrho}} \bar{Q}_{ij}^{\frac{\varrho-1}{\varrho}} \left\{ \left[P_i^{\varrho} Q_i + \sum_{j \in J_X} P_j^{\varrho} \tau_j^{1-\varrho} Q_j \right] \left[\frac{\phi(\varrho-1)}{\varrho} \right]^{\varrho} W^{-\varrho} \right\}^{\frac{\varrho-1}{\varrho} \frac{\phi}{\phi+\varrho-\varrho\phi}} (A_i)^{\frac{\varrho-1}{\phi+\varrho-\phi\varrho}} N_i^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}}, \quad j \in J_X \cap J_F,$$

$$\tilde{R}_{ij} = 0, \quad j \notin \{i\} \cup J_F \cup J_X.$$

With this definition, it's straightforward to show that the revenue share $\frac{R_j}{\sum_j R_j} = \frac{\tilde{R}_{ij}}{\sum_j \tilde{R}_{ij}}$. We can further obtain

$$\max_{z, \lambda_{TH}, \lambda_{LT}} \Bigg\{ (1-(1-\theta)\tau_i) \left[\left(1 - \lambda \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) \right) (\bar{R}_{ii} - \bar{C}_{ii}) + (1-\lambda) \sum_{j \in J_F} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) \right]$$
$$-(1-(1-\theta)\tau_i) \left[(C_{i,LT}(\lambda_{LT}) + C_{i,TH}(\lambda_{TH}) - \varphi_{iLT}\lambda_{LT} - \varphi_{iTH}\lambda_{TH}) \sum_{j \in J_F \cup \{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) \right]$$
$$+(1-(1-\theta)\tau_{LT}) \Bigg[\left(1 - (1-\lambda_{LT}) \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) \right) (\bar{R}_{i,LT} - \bar{C}_{i,LT}) +$$
$$\lambda_{LT} \sum_{j \in J_F \cup \{i\} \setminus \{LT\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) \Bigg]$$
$$-(1-(1-\theta)\tau_{LT}) \left[\varphi_{iLT}\lambda_{LT} \sum_{j \in J_F \cup \{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) \right]$$
$$+(1-(1-\theta)\tau_{TH}) \left[\lambda_{TH} \sum_j \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) - \varphi_{iTH}\lambda_{TH} \sum_{j \in J_F \cup \{i\}} \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) \right]$$
$$+ \sum_{j \in J_F \setminus \{LT\}} (1-\tau_j) \left[\left(1 - \frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) \right] \Bigg\} z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} - (1-(1-\theta)\tau_i) W_i z / A_i$$
$$- \sum_{j \in \{i\} \cup J_X \cup J_F} \tau_j \theta \cdot \frac{\tilde{R}_{ij}}{\sum_j \tilde{R}_{ij}} \cdot \Bigg\{ \sum_k \left(\bar{R}_{ik} - \bar{C}_{ik} \right) z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} - W_i z / A_i -$$
$$C(\lambda) \cdot \sum_j \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho} \right) (\bar{R}_{ij} - \bar{C}_{ij}) z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}} \Bigg\}.$$

As before, the shift shares λ_{TH} and λ_{LT} can be solved independently of z:

$$\lambda_{TH} = \left(\mathcal{C}'_{i,LT}\right)^{-1}\left[(1-\varphi_{iTH})\frac{(1-\theta)\,(\tau_i-\tau_{TH})}{1-(1-\theta)\,\tau_i-\theta\sum_k \tau_k \frac{\bar{R}_{ik}}{\sum_j \bar{R}_{ij}}}\right],$$

$$\lambda_{LT} = \left(\mathcal{C}'_{i,TH}\right)^{-1}\left[(1-\varphi_{iLT})\frac{(1-\theta)\,(\tau_i-\tau_{LT})}{1-(1-\theta)\,\tau_i-\theta\sum_k \tau_k \frac{\bar{R}_{ik}}{\sum_j \bar{R}_{ij}}}\right].$$

The FOC for z is

$$\begin{aligned}
\left(1-(1-\theta)\tau_i-\theta\sum_j \tau_j\frac{\tilde{R}_{ij}}{\sum_k \tilde{R}_{ik}}\right)W_i/A_i = \left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right) z^{\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}-1}\Bigg\{ & \sum_{j\in J_F\cup\{i\}}(1-(1-\theta)\tau_j)(\bar{R}_{ij}-\bar{C}_{ij}) \\
& -\sum_{j\in J_F\cup\{i\}}(1-\theta)(\tau_i-\tau_j)\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}) \\
& +(1-\theta)(\tau_i-\tau_{LT})\lambda_{LT}\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}) \\
& +(1-\theta)(\tau_{LT}-\tau_i)\varphi_{iLT}\lambda_{LT}\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}) \\
& +(1-\theta)(\tau_i-\tau_{TH})\lambda_{TH}\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}) \\
& +(1-\theta)(\tau_{TH}-\tau_i)\varphi_{iTH}\lambda_{TH}\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}) \\
& -(1-(1-\theta)\tau_i)\left(C_{i,LT}(\lambda_{LT})+C_{i,TH}(\lambda_{TH})\right)\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij}) \\
& -\sum_{j\in\{i\}\cup J_X\cup J_F}\tau_j\theta\cdot\frac{\tilde{R}_{ij}}{\sum_k \tilde{R}_{ik}}\cdot\Bigg[\sum_{k\in J_F\cup\{i\}}(\bar{R}_{ik}-\bar{C}_{ik}) \\
& -(C_{i,TH}(\lambda_{TH})+C_{i,LT}(\lambda_{LT}))\sum_{j\in J_F\cup\{i\}}\left(\frac{\gamma(\varrho-1)}{\phi+\varrho-\phi\varrho}\right)(\bar{R}_{ij}-\bar{C}_{ij})\Bigg]\Bigg\}.
\end{aligned}$$

Thus we can solve for optimal z as:

$$z = \left\{\left(\frac{\phi+\varrho-\phi\varrho}{\gamma(\varrho-1)}\right)\left[\frac{\left(1-(1-\theta)\tau_i-\theta\sum_j \tau_j\frac{\tilde{R}_{ij}}{\sum_k \tilde{R}_{ik}}\right)W_i/A_i}{DENOM^{PR}}\right]\right\}^{\frac{\phi+\varrho-\phi\varrho}{\gamma\varrho+\phi\varrho-\gamma-\phi-\varrho}},$$

where $DENOM^{PR}$ is the stuff inside the big brackets above.

C Data sources

World Development Indicators. Data on population and output come from the World Bank's World Development Indicators database. The specific series that we use are total population (SP.POP.TOTL), GDP in current US dollars (NY.GDP.MKTP.CD), and GDP at purchasing power parity in constant 2011 international dollars (NY.GDP.MKTP.PP.KD). For each of these variables, when constructing regional aggregates, we sum across countries with a region following McGrattan and Waddle (2020), and then average over the period 2014–2017.

World Input Output Database. International goods trade data are taken from the World Input Output Database (Timmer et al., 2015). For each bilateral import relationship, we sum all intermediate inputs and final uses of goods (industries 1–23, which represent agriculture, resource extraction, and manufacturing) from countries in the source region by countries in the destination region. We use data from 2014, the last year available.

OECD AMNE Database. This is a new data set provided by the OECD which distinguishes between three types of firms: foreign affiliates (firms with at least 50% foreign ownership), domestic MNEs (domestic firms with foreign affiliates) and domestic firms not involved in international investment. It includes a full matrix of the output of foreign affiliates in 59 countries plus the rest of the world (in the host country, industry, parent country dimension), as well as matrices for value-added, exports and imports of intermediate inputs (host country and industry). A second set of matrices in the database provides information on output, value-added, exports and imports of intermediate inputs of domestic MNEs and non-MNE domestic firms (from 2008 onwards). In addition, split Inter-Country Input-Output tables are provided distinguishing for all countries the transactions of domestic-owned and foreign-owned firms. These tables can be used to analyse multinational production in value-added terms. We exploit them to discipline our model and make sure it replicates the share of each region's gross value added that is accounted for by foreign multinationals. We first map the set of 59 countries from AMNE data set to our 5 regions and then compute the average value added shares for three types of firms (foreign affiliates, domestic MNES and domestic non-MNEs) in each region over the time period 2008-2016. The data can be accessed at OECD AMNE Database.

Compustat. Data on sales, employment, and country of origin of parent companies comes from the Compustat North America Fundamentals Annual database. This database contains data by North American companies parsed from SEC filings. Data on subsidiaries comes from the Wharton Research and Data Services (WRDS) Subsidiary Data. This data also comes from SEC filing, particularly Exhibit 21, in which firms filing with the SEC must list the names of all existing Significant Subsidiaries. A detailed, legal definition of Significant Subsidiaries, see here. Roughly, if the parent company controls at least 10% of the subsidiary, it is considered Significant. The data is available between 1995-2019, and contains identifying information for the parent company, as well as the name and country of residence of all Significant Subsidiaries. These two datasets were linked using a common identifier of the parent company, the gvkey. Mean and median sales and employment statistics were computed for years 2010-2019. The unit of observation was a parent company—year.

U.S. Census Data. To discipline the firm size distribution we exploit data from the Statistics of U.S. Businesses (SUSB). SUSB is an annual series that provides national and subnational data on the distribution of economic data by establishment industry and enterprise size. SUSB covers most of the country's economic activity. The series excludes data on nonemployer businesses, private households, railroads, agricultural production, and most government entities. We construct a Lorenz employment curve for the U.S. at the firm level using two Excel spreadsheets available at the Census website. We combine the table with detailed employment sizes with the table with larger employment sizes (20,000+ employees), both from 2019 SUSB. This allows us to account for a long, right tail of the firm size distribution in our model, which is crucial given that average MNE is three orders of magnitude larger than the average firm in the U.S. economy. Both Excel files are downloaded from the SUSB website.

Tørsløv et al. (2022). Two kinds of data are taken from this paper: lost profits and effective corporate income tax rates. Total lost profits are from sheet Table3 of the Main Data Excel file. We first sum across all countries within the North America and Europe regions, and then set the rest of the world's lost profits by subtracting the North America and Europe totals from the overall world total. The share of lost profits that are shifted to the tax haven region is constructed in the same way using sheet TableC2 in the Replication Guide Tables Excel file. The effective corporate income tax rates come from sheet DataF2 in the Main Tables Excel file. Here, we take the average across countries within each region. Both Excel files are downloaded from `https://missingprofits.world/`.

D Elasticity of profit shifting margin

In this section we discuss briefly the empirical literature on profit shifting, which aims to estimate the elasticity of reported profits with respect to the tax rate differentials across jurisdictions. We begin with an overview of the empirical strategy adapted in this line of research, then move to the discussion of the headline, consensus estimates emerging from the literature and finally we link our structural modelling approach with it.

D.1 Empirical strategy

Most of the empirical literature on elasticity of the profit shifting margin follows the concept presented by Grubert and Mutti (1991) and Hines and Rice (1994) that the reported pre-tax profit of a multinational entity, Π_i^R, is a sum of the "true" profit, Π_i^T, and the profit shifted for tax reasons, Π_i^S i.e.:

$$\Pi_i^R = \Pi_i^T + \Pi_i^S \tag{99}$$

This shifted profit would be positive in low-tax countries and negative in high-tax countries. The idea here is that a actual profitability of the multinational enterprises with alike characteristics (e.g. size, industry, country etc.) is similar. However, the opportunities to shift profits differ since they depend on such characteristics as locations of the other subsidiaries and statutory tax rates in these locations. Thus, the entities linked to low-tax jurisdictions are more likely to shift profits and the entities linked to high-tax jurisdictions are more likely to receive profits. The fundamental challenge with estimating the elasticity of profit shifting margin is that neither "true" profits nor shifted ones are directly observable in the firm-level data. To tackle this problem the literature usually assumes that "true" profits are equal to output minus the wage bill, with the wage being equal to marginal product of labor (see for example Huizinga and Laeven (2008)). As for the shifted profits the literature typically specifies some stylized framework which allows linking shifted profits to tax differential between jurisdiction j and other jurisdictions operate. This strategy leads to the following generic equation estimated to identify shifting profits:

$$\pi^R_{i,j,t} = \beta X_{i,j,t} - \gamma C_{i,j,t} + \delta_t + \varepsilon \tag{100}$$

where $\pi^R_{i,j,t} = \ln \Pi^R_{i,j,t}$ are logged, reported profits of a multinational i located in jurisdiction j at time t, $X_{i,j,t}$ is a vector of determinants of true profitability, which includes among others capital and labor inputs. It may also include a number of macroeconomic variables, such as GDP growth, exchange rate or inflation. $C_{i,j,t}$ is a composite variable that summarizes the tax differentials between jurisdiction j and other jurisdictions that the MNE located in jurisdiction i has subsidiaries in. The specific formula for $C_{i,j,t}$ differs across papers but in all of them it reflects the tax incentives to shift profits away/in from/to jurisdiction j. Finally δ_t denotes time fixed-effect and ε denotes the residual term. The coefficient of interest is then γ which reflects the extent to which the multinational shifts profits into or out of affiliate i. It is important to note that this estimate represents a marginal effect – i.e. the change in reported profits associated with a small change in tax rates, *holding all else constant.*We can interpret γ in equation 100 was the semi-elasticity of observed profits π^O_i with respect to the composite tax variable $C_{i,j,t}$. The *semi-elasticity* indicates the percentage change of reported profit in response to a one percentage point change in the tax differential vis-'a-vis other international locations, reflecting the incentive to shift profits abroad, i.e.:

$$\text{Semi-Elasticity} = \frac{\partial ln \text{ Reported Profits}}{\partial \text{ Tax Differential}} \approx \frac{\partial \text{ Reported Profit}}{\text{Reported Profit}} \times \frac{1}{\partial \text{ Tax Differential}}$$

Note that differentiating 100 we get:

$$\frac{\partial \pi^R_{i,j,t}}{\partial C_{i,j,t}} = -\frac{\partial \Pi^R_{i,j,t}}{\Pi^R_{i,j,t}} \frac{1}{\partial C_{i,j,t}} = \gamma \tag{101}$$

thus γ reflects indeed the semi-elasticity of interest.

D.2 Empirical estimates

A number of papers estimate different versions of equation 100 for variety of data sets and time periods. A thorough and detailed review of this literature is beyond the scope of this paper.[21] Instead, we focus here on the two most recent survey papers, that conduct a meta analysis of existing estimates, and on the main OECD estimate, all reporting the headline, semi-elasticity number.

Johansson et al. (2017) provide the main estimate of the magnitude of the profit shifting used by the OECD. They conduct a comprehensive study using firm-level data from the ORBIS database to assess international tax planning by multinational enterprises (MNEs). Their results are based on an impressive, very large sample of firms (1.2 million observations of MNE accounts) in 46 OECD and G20 countries and a sophisticated procedure to identify MNE groups. Their headline estimate of the semi-elasticity of the profit shifting margin with respect to the tax differential is **1.11** (see Table 1, column 1 and footnote 31 in their paper). Hence, reported profits decrease by about 1.1% if the international tax rate differential increases by one percentage point. The estimated elasticities combined with a number of assumptions are then used to estimate the effect of international tax planning on corporate tax revenues: the estimated net tax revenue

[21]See Dharmapala (2014), Heckemeyer and Overesch (2017), Johansson et al. (2017) and Beer et al. (2020) for extensive reviews of this line of research.

loss ranges from 4% to 10% of global corporate tax revenues.

Heckemeyer and Overesch (2017) construct a meta-database containing 203 primary estimates sampled from 27 empirical studies identified by means of article search engines. All of the included studies estimate the empirical relationship between reported parent and subsidiary profitability and the tax incentive to shift profits abroad. Therefore, this meta-analysis reviews the literature, which provides indirect evidence for profit shifting without specifying directly the shifting methods. They find a tax semi-elasticity of pre-tax profit of about **0.79**, in absolute terms. They conclude that across all specifications the predicted semi-elasticities turn out statistically significant and rather robust in magnitude. They also provide a 95% confidence interval in addition to the point estimate and conclude that conditional on a hypothetical state-of-the-art study design, the set of semi-elasticities that we they would not reject at the 5% significance level ranges from 0.546 to 1.026.

Beer, de Mooij, and Liu (2020) extend the analysis conducted by Heckemeyer and Overesch (2017) and include 11 additional studies and 199 additional primary estimates. They also reduce specification bias, and adopt an enhanced estimation method that corrects for within-study correlation of primary estimates. Their results indicate that a semielasticity of reported pretax profits with respect to international tax differentials equal to **0.98** is a good reflection of the literature. This means that a 1 percentage-point larger tax rate differential reduces reported pretax profits of an affiliate by 1%.

D.3 Model counterpart of semi-elasticity

We now describe how we estimate the model counterpart of the semi-elasticity summarized above. We view this as a validation exercise of the cost function $C(\lambda)$ upon which the extent of profit shifting in the presence of tax differentials between jurisdictions heavily depends. Since our parsimonious model of only four productive regions does not provide sufficient variation in cross-jurisdiction differences in corporate tax rates (regressor $C_{i,j,t}$ in equation 100), we conduct a simulation exercise as follows.

We simulate 100 counterfactual economies, of which the first 50 economies we raise the corporate tax rate of the LT region incrementally and latter 50 we raise the rate of the TH region. We set the highest counterfactual corporate tax rate to 15%, equal to the global minimum tax rate of OECD Pillar 2. In each of these counterfactual economies, we hold fix the set of firms' FDI and exporting destinations, J_F and J_X, and in addition the final good price and wage rate of each region, P_i and W_i. We allow firms to solve for their optimal choices of labor $\boldsymbol{\ell}$, intangible capital z and shifting share λ_{LT} and λ_{TH}. In other words, the firms' problem is re-sovled in a partial equilibrium fashion, which allows us to isolate the relationship of reported profits in home divisions to tax rate differentials relative to the profit shifting destination. Denote k as the index of a counterfactual economy.

We follow the empirical specification of equation 100 and run the regression using the model-simulated dataset:

$$\log \pi_i^{k,PS}(\omega) = \beta_0 + \beta_\ell \log \ell_i^k(\omega) + \beta_z \log z^k(\omega) + \beta_\tau \hat{\tau}_i^k + \epsilon_i^k(\omega) \tag{102}$$

where we denote τ_i^k as the counterfactual tax differential defined as $\hat{\tau}_i^k = \tau_i - \tau_{LT}^k$ for $k \leq 50$ and $\hat{\tau}_i^k = \tau_i - \tau_{TH}^k$ otherwise. For each experiment k, we include in the regression only home divisions of firms doing FDI in the region of which we change the corporate tax rate. We only include home divisions of profit-shifting MNEs because we do not model profit shifting originating from a foreign subsidiary. Nonetheless, such regression informs us of how reported profit responds to changes in profit shifting relevant tax differentials, which is captured by the coefficient of interest β_τ. We report the coefficient estimate of β_τ in Table 3.