Utrecht University

Economic Policy 2021/2022

Master Thesis

Harmen de Jong

5732913

Harmen.dejong@planet.nl

How high should the Minimum Tax be?

A theoretical analysis of the OECD's proposed global minimum corporate tax structure

Content

(10169 words, excluding footnotes, references and appendices)

1. Introduction

1.1. The OECD/G20 Inclusive Framework on BEPS

In 2016, the Organisation for Economic Co-operation and Development (OECD) created the Inclusive Framework on BEPS. This multilateral organization seeks to create and implement policy responses to economic developments that rendered existing global tax agreements outdated. National economies are increasingly digitalized and integrated, allowing multinational enterprises (MNE)s to engage in base erosion profit shifting. This is a strategy where MNEs minimize their tax burden by declaring profits in locations with low tax rates or by exploiting mismatches in tax regulations between jurisdictions. For example, First Quantum Minerals is a Canadian mining company. For mining operations in Finland, it set up a Swedish holding company between the Canadian parent company and its Finnish subsidiaries to minimize its tax liabilities by exploiting a discrepancy between Finnish and Swedish tax law (Finer & Ylönen, 2017). BEPS is estimated to cause an annual tax revenue loss of 100 to 240 billion USD worldwide, equivalent to 4% to 10% of global corporate income tax revenue (OECD, 2017).

The Inclusive Framework published blueprints for a global plan to combat profit shifting in 2020. The plan consists of two pillars. Pillar One creates global standards for tax allocation rules. It aims to standardize what is taxed and which jurisdiction has the right to tax it. Notably, it gives a part of the taxing rights to the jurisdiction where the revenue is sourced (OECD, 2020a). This eliminates inconsistencies between national tax codes that MNEs can exploit as loopholes. Pillar Two establishes a global minimum corporate tax rate. The tax is implemented as follows: the proposal gives a formulaic definition for an MNE's Effective Tax Rate (ETR). If that rate is lower than the minimum rate, governments have the right to charge that MNE additional taxes equal to the amount required to make the MNE's ETR equal to the minimum rate (OECD, 2020b; OECD, 2021).

1.2. Research approach

This paper concerns the proposals for a global minimum tax in Pillar Two. Particularly the level at which the minimum rate should be set. This rate is set at 15% in a recent version of Pillar Two(OECD 2021) however, this rate is subject to discussion. Pillar Two was never designed with a specific level of minimum tax in mind and the design leaves this up to the preferences of policymakers instead. For example, the OECD's impact assessment, which estimates global revenue gain from Pillar Two, provides estimates for a range of possible minimum rates, from 7.5% to 17.5% (OECD 2020c). The cover statement in the blueprints of Pillar Two also states that the minimum rate should be 'an agreed rate' (OECD 2020b).

Yet the height of the minimum rate is more than a simple matter of taste; it can greatly affect the functionality or efficiency of the entire design. In economic literature, most criticism of Pillar Two is either directly targeted at the height of the minimum rate or rests upon a mechanism that is heavily contingent on the height of the minimum rate, mostly because a sufficiently high rate would break down tax enforcement cooperation or would impede on the ability of developing nations to attract foreign direct investments. I elaborate on this in chapter 2.

The problem here is that the height of the minimum rate is left open-ended by design despite the particular rate being chosen having serious implications for the mechanism of Pillar Two and its impact on the world. In this paper, I aim to address that problem. I calculate how tax revenues per jurisdiction change as a consequence of the introduction of a minimum corporate income tax rate (MR), using data on profit shifting from Tørsløv et al. (2018) and mathematical modelling of the relationship between tax rates and profit shifting based on the Kanbur-Keen model (Kanbur and Keen 1993; Keen and Konrad

2013). Using this output I recommend an optimal minimum rate for various metrics of optimality. Notably, this method excludes real economic activity of MNEs and only captures changes in tax revenue caused by a change in a tax jurisdiction's tax rate or by a change in profit shifting. This specification keeps my research close to the target of reducing profit shifting, rather than embedding it in a broader debate about the value of public goods compared to economic losses from taxation. This is also beneficial for modelling actor behaviour.[1] A model within these limitations should be capable of addressing the concerns of a minimum rate found in the literature.[2]

The main questions I seek to answer are: within which range of rates should the MR be placed? And which rate is likely optimal? Hebous and Keen (2022) have a good metric for an advisable range; the range between the Pareto efficient rate and the maximum Pareto dominant rate. The Pareto efficient minimum rate is the rate for which no jurisdiction can be made better off without making at least one jurisdiction worse off compared to the previous rate. Logically, this should be identical to the optimal rate for havens, as non-havens will always benefit from an increase in the minimum rate. The maximum Pareto dominant rate is the best possible rate for non-havens for which no jurisdiction is worse off compared to the world before Pillar Two.[3] Pareto improving rates make good minimum rates because they avoid making certain jurisdictions worse off. This is important for getting widespread international support for the reform. The best possible rate that is Pareto improving compared to the world before the reform must logically be within the range between the Pareto efficient rate and the maximum Pareto dominant rate, as the former is more beneficial to both havens and non-havens than any rate lower than it and any rate higher than the latter is not Pareto improving compared to the pre-Pillar Two world. In addition to Pareto improving rates, I will include figures for welfare gain, as that is the most straightforward, commonly used metric of optimality in economics in general.

1.3. Societal relevance of research

Tax structure is of great relevance for public debate. Tax is the primary source of revenue for most governments and a major burden for individuals and businesses. An efficient tax system that makes that trade-off in the best possible way is therefore extremely important in maintaining welfare in society and benefits all parties greatly. Thus, anything that increases the efficiency of the tax system benefits society to a large extent.

This is particularly true for finding a range of advisable rates for the minimum rate in Pillar Two. As the OECD points out, Pillar Two has the potential to be a huge revenue gain and partially eliminate profit shifting, one of the largest inefficiencies in the global corporate tax system. However, as explained in chapter 2.1., the literature suggests that an incorrectly set minimum rate can have adverse effects that would prevent this welfare gain and could even make problems related to profit shifting worse. Therefore, my research question directly addresses a global policy problem. Additionally, that policy problem is still unsolved. So far, no study has decisively proven how high its MR should be or whether or not the current proposal of a MR of 15% is in the range of prudent rates. Therefore, knowledge on how to set the minimum rate efficiently is crucial in maximizing the benefits of Pillar Two.

Currently the world is debating whether or not to implement the reform proposed by the Inclusive Framework. The EU has not unanimously agreed it yet, because Hungary is still opposed. In the United States it is yet unclear if there will be a Senate majority in favour of legal changes required to implement Pillar Two. The outcome of the 2022 midterm election can be decisive in whether or not

[1] See chapter 4.3.
[2] See chapter 2.1.
[3] See chapter 4.1. for a straightforward model of these concepts

the US will support Pillar Two. Knowing how to set a minimum rate that minimizes the amount of losing parties can greatly help the quality of the debate. Parties that stand to lose have an incentive to oppose any reform, therefore the world is more likely to find agreement on something if fewer parties are losing.

2. Literature review

2.1. Debate around the Minimum Rate

Academic literature on Pillar Two usually fits in one of two strands. The first is the extent to which Pillar Two is compatible with pre-existing law and legal principles, notably tax sovereignty of individual jurisdictions. This is covered in Massant and De Broe (2021), Brokelind (2021) and Da Silva (2020), among others. The second strand is the discussion about whether or not the Pillar Two design will be effective at combatting profit shifting and providing governments with a more optimal tax system. For the purpose of this paper, I am interested in the latter.

Johannesen (2022) finds that a high level of minimum tax is necessary to ensure a positive net welfare effect of minimum taxes and that it is unclear if 15% is enough. This is because profit shifting will continue to happen if the global minimum tax is lower than the corporate tax rate of non-tax-havens. In that scenario, tax havens are incentivized to raise their corporate tax rate to the global minimum tax rate to ensure that the tax is paid in their jurisdiction. Since profit shifting still exists, though reduced in scope, a minimum tax has two opposing welfare effects on non-havens. A positive effect caused by the reduction in profit shifting and a negative effect caused by increased corporate tax rates in tax havens. The latter is negative, since economic activities of MNEs primarily take place in non-tax havens, meaning that an increase in taxes the MNE has to pay to havens will cost the MNE resources that it would otherwise spend in non-havens.

Hindriks and Nishimura (2022) find that a high level of minimum tax can have adverse effects. They argue that the introduction of a global minimum tax reduces the incentive of a tax haven to cooperate globally on tax enforcement, which decreases the costs of profit shifting and thus increases incentive for MNEs to engage in profit shifting. This leads to a welfare loss in non-tax-havens and consequentially, the authors argue that enforcement strategy needs to be considered more thoroughly in Pillar Two. As explained in chapter 3.1, the mechanisms of Pillar Two require data on MNE group entities in low-tax jurisdictions. if low-tax jurisdiction refuse to collect or share necessary data on MNEs, this can severely impede Pillar Two's capacity to monitor and reduce profit shifting. In the Hindriks-Nishimura model, tax cooperation breaks down when the welfare loss on the low-tax nation reaches a certain threshold where the costs of complying with the tax agreement would outweigh its benefits. Whether or not this happens depends on how high the minimum rate is set.

The views expressed by Johannesen and Hindriks and Nishimura seem contradictory, but they are not mutually exclusive. Rather, they emphasize how important it is to make the correct trade-off in determining the MR. If it is too low, it hardly reduces incentive for MNEs to shift their profits and the main difference it makes is that tax havens can increase their tax rates without fear of competition from other havens with even lower rates. This leads to a situation like Johannessen warns of, where almost all of the revenue gained from Pillar Two accrues in tax havens and non-havens face the economic effects caused by their firms facing higher costs. In this situation, international tax rules heavily incentivize jurisdictions to be tax havens, meaning that more will choose to become one and profit shifting is likely to get worse. If the MR is too high, it will reduce profit shifting to such a large extent that the minimum tax will cost tax havens much more than they gain from it. At that point, they are incentivized to undermine the system. Since the system is vulnerable to the consequences of non-cooperation, this leads to the welfare loss Hindriks and Nishimura warn for.

Both of these papers are successful at demonstrating the adverse consequences of respectively too low and too high minimum rates. This discourse still lacks a way to make the trade-off between benefiting havens too much or too little. My paper aims to fill in that gap. The model I specified captures the effect from both papers. The incentives of tax havens in Hindriks and Nishimura are all expressed in terms of revenue gain or loss. Although the adverse effect for non-havens in Johannesen comes in the form of loss in real economic activity, the rates at which this occurs are identifiable by excessive revenue gain in tax havens, which is included in the model.

2.2. Other models of Pillar Two

The model in this paper is not the first model of the effects of Pillar Two. To further justify the added value of my research to the literature, I briefly go over other models and explain what my model uniquely adds in comparison to the other models and why it is better suited to answer my research question and solve the dilemma presented in 2.1.

The OECD's economists find that, assuming a minimum rate of 15%, Pillar Two would increase corporate income tax revenue around the world by 1.8 to 4.2%. This is a consequence of reduced profit shifting and increased tax rates. They concede that Pillar Two would increase the average investment costs of MNEs, which reduces global investments. However, they argue that this reduction is very small, because Pillar Two mostly affects MNEs with high profit margins, whose investment decisions are relatively less sensitive to tax levels. Additionally, they argue that this effect is offset by implicit positive effects of Pillar Two on investment and economic growth. First, reducing tax cooperation makes non-tax factors more important in investment decisions, which leads to a more efficient global capital allocation. Second, the additional tax revenue reduces the need for governments to levy other taxes, that are potentially more distortive (OECD, 2020c). There are two problems with relying on this model to determine optimal rates. The first is that for the sake of transparency it is better to have an independent actor estimate the proposals' impact. The second is that the OECD estimates impact on a global level, whereas to make the trade-off discussed in 2.1., estimates on a national level are needed.

The World Investment Report 2022 by the United Nations modelled the effects of Pillar Two on investment. This is because Foreign Direct Investments (FDI) are crucial for developing economies, since they need to attract capital to grow, and thus any distortion of this is very important in developmental economics. The report estimates that Pillar Two will reduce the global volume of FDI by 1-4%, depending on which assumptions are applied. However, Pillar Two also affects the distribution of FDI, and this is an area where developing nations are expected to gain. The report estimates that 'Offshore Financial Centers' (tax havens) will lose a disproportionate 4.0 to 7.3% of their FDI and that developing economies will be the biggest beneficiaries of that shift (UNCTAD 2022). While this can have important consequences for the implementation of Pillar Two, further discussed in 6.2, it models a different aspect of minimum taxation than needed in this context.

Hebous and Keen (2022) create a very similar model to mine. They are also interested in the height of the minimum rate and base their calculations on the Kanbur-Keen model. They find that both tax havens and non-havens can benefit substantially from a high minimum rate. For example, when using Ireland as a representative for havens, they find a Pareto efficient MR of 18.75% and a Pareto dominant rate of 25%. However, there are two issues with these estimates. One is that their model is largely theoretical and uses very few real-world inputs. Such a model is good at showing how the mechanics of Pillar Two work, but less good at determining specific values, such as the height of the MR, as those depend heavily on model inputs. Two, their model bases tax rates in havens on the statutory tax rate in Ireland, 12.5%. This rate is substantially higher than effective tax rates. The effective tax rate in tax havens, weighted by the amount of profit shifted to each haven, is 5.8%.[4] This discrepancy likely explains why their estimates are so high. A model that uses effective tax rates should be preferred, as those are a better indicator of how much tax an MNE actually pays on shifted profits.

Finally, the Oxford University Centre for Business Taxation released an evaluation of Pillar Two (Devereaux et al. 2020). They evaluate Pillar Two from both legal and economic perspectives and make recommendations based on that. The authors present similar concerns to Hindriks and Nishimura (2022), arguing that Pillar Two will only be effective if the proposals are adopted almost universally. They add to this that harmonization of accounting standards is necessary to make the proposals implementable. Their paper contains estimates for revenue gain on a jurisdictional basis, however those are based on the assumption that there is no change in amount of profit shifting. In a later section it does contain a model for changes in profit shifting, but that is based on arbitrary inputs. Both of these serve the purposes of the authors, but cannot be used to answer my research question.

3. Theoretical framework

3.1 Regulatory framework

The Pillar Two framework goes beyond simply mandating jurisdictions to change their nominal tax rate. It uses complex mechanisms to achieve its goal of taxing MNEs a percentage of their profit that is at least equal to the minimum rate. Some consideration of how these mechanisms work is needed to make the model in this paper match the system of Pillar Two.

The rules in Pillar Two are called the GloBE rules. The GloBE rules apply to MNE groups with an annual revenue of 750 million euro or above. MNE groups are defined as any groups of Entities that are related through ownership or control which include at least one Entity or Permanent Establishment not located in the jurisdiction of the Ultimate Parent Entity. An Ultimate Parent Entity is either an Entity that directly or indirectly owns a Controlling Interest in any other Entity and is not owned by any other entity or an entity that is not part of any group but does own permanent establishments in other jurisdictions. Government entities, Non-Profit organizations and pension funds are excluded from the GloBE rules (OECD 2021).

The Income Inclusion Rule is the main mechanism through which the minimum rate materializes in Pillar Two. In short, the rule states that if an entity in a MNE group has an effective tax rate that is lower than the minimum rate, then the Ultimate Parent Entity of that group has to pay the percentage of that entity's profit equal to the MR minus its current ETR.[5] The jurisdiction in which the Ultimate Parent Entity is located has the right to levy this tax (OECD 2021).

If the jurisdiction of the Ultimate Parent Entity chooses not to levy this tax, the GloBE rules have a secondary mechanism to enforce its minimum rate, the Undertaxed Payments Rule. Under this rule, the right to tax the MNE group the due amount is divided among jurisdictions that have implemented

[4] The difference between statutory and effective tax rate in relation to the Pillar Two proposals is explained in chapter 3.1. The 5.8% comes from the calculation in chapter 4.4.

[5] To account for shared ownership, the due amount is finally multiplied by this ratio: $\frac{(A-B)}{A}$, where A is the income of the entity in question for the fiscal year in question and B is the amount of that income which is attributable to other owners than the Ultimate Parent Entity.

the UTPR and have an entity from the MNE group located in them. The share of the amount of due revenue allocated to each jurisdiction is for 50% determined by the share of the MNE group's employees residing in that jurisdiction and 50% by the share of the value of tangible assets located in that jurisdiction.

The GloBE rules use effective tax rates (ETR) instead of statutory tax rates. The ETR is calculated on a jurisdictional basis. This means ETR is the total amount of income taxes paid in a jurisdiction for a fiscal year divided by the total profits of the entities within that jurisdiction for that same year. The statutory tax rate is the legally imposed tax rate. This is a different rate than ETR because tax jurisdictions typically have laws that allow corporate entities to deduct certain things from their taxes, meaning that the actual percentage of taxes paid - the ETR - is lower than the statutory tax rate. I use ETR instead of statutory tax rate in my model, because the GloBE rules use ETR and because ETR is a better indicator of how much taxes are actually paid. Because using the GloBE rules requires the calculation of the ETR for each entity of an MNE group, Pillar Two requires all jurisdictions to share data on profits booked and taxes paid in their jurisdiction. This makes the proposals vulnerable to the non-compliance problem Hindriks and Nishimura (2022) outline.

3.2 The Kanbur-Keen model

The model in this study is based on the Kanbur-Keen model of tax competition, developed by Kanbur and Keen (1993). This model is the best-suited model for a study of the GloBE rules for the following reason: The Kanbur-Keen model does not assume that the tax base that countries compete over is mechanically tied to real economic activity, unlike its most widely used counterpart, the Zodrow-Mieszkowski-Wilson model, developed by Zodrow and Mieszkowski (1986) and Wilson (1986). Instead, the Kanbur-Keen model features a design for modelling profit shifting (Keen and Konrad 2013).

The Kanbur-Keen model was originally designed to model tax competition in value-added tax (VAT), but as one of the authors noted in a later work, the same logic can be applied to corporate tax competition and profit shifting (Keen and Konrad 2013). In the Kanbur-Keen model, there are two jurisdictions, one with high VAT rates and one with low VAT rates. All households in both jurisdictions want to consume one unit of consumable goods with a price of 1. This means that in the highly taxed nation, consumers have to decide on whether to shop domestically, or go abroad and buy the unit at a lower tax rate. This means that tax revenues are:

$$T_{Low} = \tau_{Low}\left(h_{Low} + h_{High} * s\right); \quad T_{High} = \tau_{High} h_{High}(1 - s) \tag{1}$$

Where T is tax revenue, τ is the tax rate, h is the amount of households[6] living in each nation, 'High' is the highly taxed nation, 'Low' is the lowly taxed nation and *s* is the proportion of households in High that choose to buy in Low. Because tax revenue is the tax rate multiplied by the tax base, equation (1) finds tax revenues for both jurisdictions by subtracting the consumers from High that buy in Low from the tax base of High[7] and adding them to the tax base of Low.

Consumers in High face a trade-off between paying lower tax and paying the transportation costs, δ, of going to Low. The model assumes customers choose the outcome that gives them the most value. Because each consumer buys one unit of goods, costing one unit of currency, the difference in tax levels is equal to the amount of money each consumer saves. Thus, in equilibrium, *s* is found by:

$$s^* = \frac{(\tau_{High} - \tau_{low})}{\delta} \tag{2}$$

[6] Note that each household consumes one unit, so h is also equal to amounts of units consumed

[7] Note that $h_{High}(1 - s) = h_{High} - h_{High} * s$

A logical conclusion of this model is that smaller countries benefit more from lower tax rates than larger countries, since the absolute amount of tax revenue raised by domestic firms is relatively small, it is comparatively advantageous to them to aim to attract foreign firms to their jurisdiction. Consequently, it means that setting a uniform global corporate tax rate cannot be advantageous to both parties, as both countries have different optimal tax rates. A global minimum tax, however, can be advantageous if the rate is within a certain range (Kanbur and Keen, 1993; Keen and Konrad 2013).

3.3 Application to profit shifting

To apply the model to profit shifting, only minor changes have to be made, since the logic is largely the same. Firms in High need to decide between declaring profits domestically or abroad. The latter comes at a lower tax rate but comes at a 'transportation cost', namely the cost of accounting and concealment operations required to undertake profit shifting. While profit shifting erodes the tax base in high analogously to the original design, profit shifting is not real economic activity, meaning that there is no equivalent to domestic consumption in Low in a model for profit shifting. Thus, tax revenue is found by:

$$T_{Low} = \tau_{Low} * S \; ; \quad T_{High} = \tau_{High}(\Pi - S) \tag{3}$$

Where capital S is the total amount of profits shifted, and Π is the total amount of pre-tax profit generated by firms in High. Similar to the original design, firms only have an incentive to shift profits if the costs of profit shifting are smaller than the gains from profit shifting, meaning that cost can be calculated analogously, meaning[8] that:

$$S^* = \Pi * \frac{(\tau_{High} - \tau_{low})}{\delta} \; ; \; s^* = \frac{(\tau_{High} - \tau_{low})}{\delta} \tag{4}$$

3.4 Calculating costs of profit shifting

Tax rates and pre-tax profits are given. Therefore, the model can find S if δ is known. In the literature, there are two approaches to calculating δ, a theoretical and an empirical approach. The theoretical approach is an extension of the Kanbur-Keen model, found for example in Keen and Konrad (2013), Hindriks and Nishimura (2022) and Huizinga and Laeven (2008). Here, cost functions typically take the form of:

$$C = a\delta s^2 \Pi \tag{5}$$

Where C is total cost[9] and a is a parameter, usually between 1 and 0. In Keen and Konrad (2013) it has a value of 0.5, because if a is 0.5, then the partial derivative with respect to S of a function for firms' post-tax profits has equation 4 as a maximum. Notably, s^2 suggests costs are convex, yet δ is a constant, which suggests that despite that, the relationship between S^* and the tax differential is linear. In Hindriks and Nishimura (2022) δ is not quite constant, but is a function of tax enforcement efforts, meaning that δ becomes smaller if cooperation between the haven and non-haven breaks down.

Using equation 5 to calculate costs in my model is problematic for two reasons. One, C is not a function of the tax differential, meaning that it is poorly equipped to measure a response to a MR. Two,

[8] The multiplication by Π is added because S represent amount of profits shifted, not proportion of profit shifted, $S = s * \Pi$, which also explains a difference in form between equations 1 and 3. Mathematically, it makes no difference whether to calculate it with S or s, but constructing the model like this makes it easier to compile statistics on changes in the extent of profit shifting.

[9] The difference between δ and C is that δ is costs per s^2, not total costs

equation 5 contains two unknown variables, *s* and *C*, both being a function of each other.[10] Therefore, I use a variation of equation 4 from the empirical literature to calculate shifted profits.

In the empirical approach to calculating costs of profit shifting, something very similar happens. Many empirical studies into corporate income tax are interested in finding profit shifting semi-elasticities with respect to tax rate differentials, ε, using regression analysis. Semi-elasticity in this context is the percentage increase of profit shifting for each percentage increase of the tax rate differential, $\tau_{High} - \tau_{low}$. De Mooij and Ederveen (2008) and Heckemeyer and Overesch (2017) provide extensive reviews of the literature in this field. Notably, Heckemeyer and Overesch find a consensus ε of 0.8, by running regressions on results of previous studies, which control for real economic activity.

The tax semi-elasticity has the same mathematical relationship with S as δ^{-1} has with S^* in equation 4. Proof of this is in Appendix 1.1. This means that using semi-elasticities to calculate *S* is equivalent to using equation 4 under the assumption that the new amount of profit shifting will reach the equilibrium value.

4. The model

Chapter 3 has provided the theoretical foundations for a model of profit shifting. Chapter 4 outlines the design of the model. It includes methodology, data description and assumptions behind the model.

4.1. The simplified model

Using the mathematics from the previous chapter, a simplified model can demonstrate how a minimum tax changes tax revenue. Suppose we have a world with two jurisdictions, called Haven and Non-Haven. Suppose that before the implementation of Pillar Two, both jurisdictions had a fixed ETR on corporate income in their jurisdiction, 5% in Haven and 25% in Non-Haven. The economic activity of MNEs affected by the Pillar Two proposal in Non-Haven is worth 500 billion euros. MNEs operating in Non-Haven shift profits to Haven with a tax semi-elasticity of 0.8, in accordance with the findings of Heckemeyer and Overesch (2017). Under these assumptions, the amount of profit shifted is:

$$S = \Pi * 0.8 * (\tau_{NH} - \tau_H). \quad (6)$$

The tax revenue on the MNEs income collected by Non-Haven is then:

$$T_{NH} = (\Pi - \mathrm{S}) * \tau_{NH} \quad (7)$$

The outcome of this calculation is graphically represented in red in figure 1. The graph starts flat, because if the MR is smaller than or equal to the tax rate in Non-Haven, the ETR in Non-Haven remains constant, meaning the tax differential remains constant and consequently, the amount of profit shifted remains constant. When the MR is larger than 5%, tax revenue in Non-Haven increases because less

[10] Substituting equation 4 in for *s* does not work because *C* is not constant for each *S*, meaning that pre-reform *C* cannot be used to calculate post-reform profit shifting.

profit gets shifted out of its jurisdiction. This increase is linear because we assume a fixed tax semi-elasticity of 0.8, which makes the relationship between MR and tax revenue in Non-Haven linear.

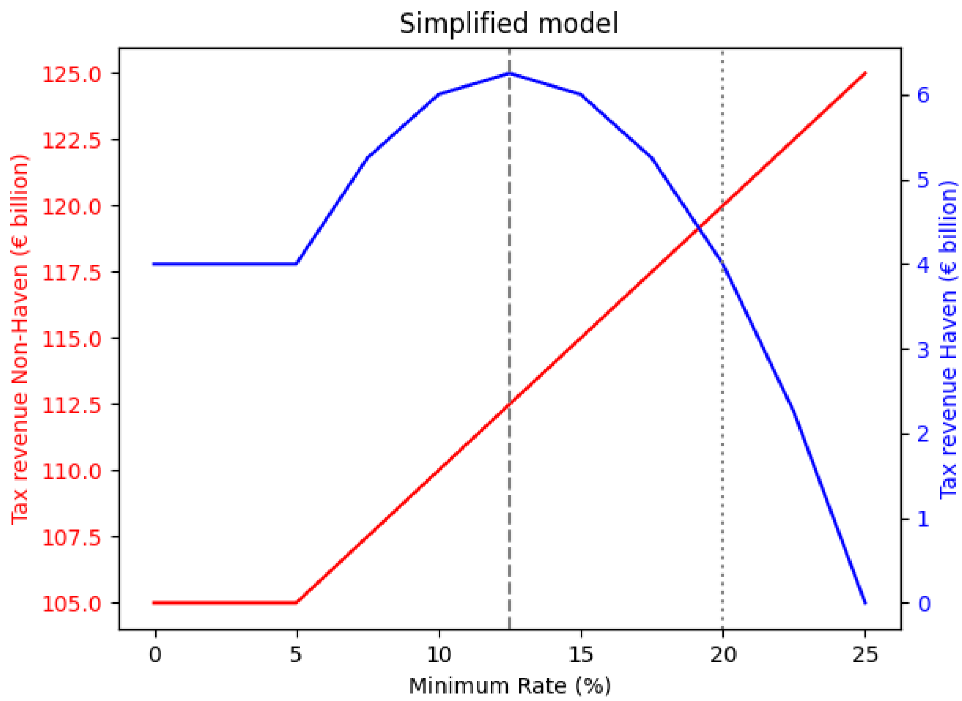

Figure 1: The simplified model. The Haven receives the largest possible amount of tax revenue at 12.5%, Revenue in haven at 20% is equal to revenue in Haven before the reform.

The tax revenue collected by Haven on profits shifted from Non-Haven is:

$$T_H = S * \tau_H \quad (8)$$

Where τ is tax rate after Pillar Two, under the assumption that Haven sets its ETR equal to the MR in Pillar Two. This is represented in Figure 2. Although the calculation looks simpler, the relationship between tax revenue and minimum rate is more complex than in Figure 1. There is no longer a simple linear relationship because an increase in MR has two converse effects on tax revenue in Haven. While a higher tax rate means that a higher percentage of shifted profits ends up as tax revenue in Haven, it also means less profit gets shifted to Haven. In this model, that relationship is parabolic because $S = \Pi * 0.8 * (\tau_{NH} - \tau_H)$, where τ_{NH} and Y are constant at respectively 0.25[11] and 500, meaning that $T_H = S * \tau_H$, can be rewritten as $T_H = \left(500 * 0.8 * (0.25 - \tau_H)\right) * \tau_H = 100\tau_H * 400{\tau_H}^2$.

In this simulation, the Pareto efficient rate is 12.5%, represented by the dashed line. Any increase in the MR from there is not Pareto improving as it makes Haven worse off. The maximum Pareto dominant rate is 20%, represented by the dotted line. Any increase from there makes Haven worse off compared to the initial situation. Thus, in this simulation, the MR should be set between 12.5% and 20%. From the perspective of the Non-Haven, a rate slightly below 20% is likely ideal in practice, as it creates a more substantial welfare gain in Non-Haven, while still making Haven better off and thus willing to cooperate.

[11] Note that this assumes the MR to be lower than or equal to the ETR in Non-Haven, meaning that the following equation does not apply to an MR greater than 25%. In those cases, both jurisdictions set their ETR equal to the MR and no profits get shifted.

The simple model shows how modelling can find the best MR for a certain set of preferences. Its main problem is that its inputs are arbitrarily chosen. For example, ETRs in the real world are not necessarily 5% in havens and 25% in non-havens, yet if we change the initial tax rates, the model will recommend a different optimal tax rate. Therefore, the model needs to be adjusted and fed real-world inputs to give policy advice that fits a real-world problem. The rest of the chapter discusses how to transform this model with real inputs.

4.2. Assumptions behind the model

Every model is based on certain assumptions. This model is no different. Firstly, the model accepts the logic of the Kanbur-Keen model and the validity of the data used, which are discussed more in-depth in their respective sections. Additionally, it has to make certain assumptions about the behaviour of the actors involved, since the model predicts how Pillar Two will change profit shifting. The current section makes those assumptions explicit and briefly discusses them.

Assumption 1: actors make rational decisions. This assumption is necessary to predict responses to the MR, namely how much profit MNEs will shift and how governments will alter their tax policy. It is an assumption commonly found in economic models, and one that is commonly criticized. Particularly problematic is that if rationality is defined as maximizing monetary gains, it neglects non-pecuniary incentives of humans. If it is defined as maximizing utility, it is unfalsifiable as utility cannot be observed. See Hodgeson (2012) for a more thorough discussion of rational choice theory.

In defence of this assumption, it should be noted that the actors in this model are governments and MNEs. While individual human beings have limited time and information to make decisions, governments and MNEs have elaborate decision-making structures with countless employees devising optimal strategy for their actor. Insofar as rational choice theory is flawed, it is least flawed in this context.

Assumption 2: MNEs maximize their utility by maximizing profits. This is a more trivial assumption. MNEs are ultimately controlled by shareholders, who are primarily interested in getting a return on their investment. I am unaware of any argument against this assumption that goes beyond the nuance that individual actors within an MNE might have their own specific incentives.

Assumption 3: Governments maximize their utility by maximizing tax revenue. This assumption might appear to be equally trivial as assumption 2, but it is not. In reality, governments have a wide range of incentives, including appealing to the electorate, stimulating economic growth and meeting their diplomatic goals. Notably, in the context of corporate income tax, governments have an incentive to attract foreign direct investment by having lower tax rates than competing jurisdictions. Therefore, this assumption cannot be true in a holistic model of the effects of Pillar 2.

The problem is that there is no workable alternative to this assumption, because the utility of governments is unfindable without it, as there is no objective measure of how much governments value foreign direct investment compared to tax revenue. Most likely, this trade-off is different for each nation, and for each political faction within each nation. And without knowing this, it is impossible to calculate how high governments prefer the MR to be.

Therefore, I've chosen to model exclusively profit shifting effects of Pillar Two. In this context, assumption 3 is plausible since profit shifting is not connected to real economic activity, and when real economic activity is excluded, it stands to reason that governments prefer the rate that maximizes their revenue. Also, note that there are reasons to assume that real economic activity will not be severely affected. Firstly, because the Income Inclusion Rule and Undertaxed Payments Rule effectively eliminate tax competition below the minimum rate. Secondly, because, firms with high profitability ratios are less sensitive to tax changes than the average MNE and Pillar Two only applies to MNE groups with 750 million euro profits or above, most of which have a high profitability ratio (OECD 2020c). Thirdly, because the alternative to raising government revenue through corporate income tax is raising it through another tax that is likely to distort economic activity (OECD 2020c)

As a consequence of the mechanisms of Pillar Two and assumptions 1 and 3, governments of tax jurisdictions with an ETR lower than the MR will increase their tax rate until their ETR is equal to the MR. Because if they do not, the Income Inclusion Rule and the Undertaxed Payments Rule would allow other governments to tax profit declared in their jurisdiction, meaning that the tax MNEs have to pay on that profit would be the same as if they had charged the MR. That means that setting the tax rate below the MR costs potential tax revenue, but does not attract more investment or more shifted profits.

Assumption 4: Governments do not change their tax policy, unless their ETR is below the MR. The problem with this assumption is that a decrease in profit shifting means that losses from having high taxes decrease, which might incentivize governments to increase corporate tax rates. In fact, the Kanbur-Keen model provides response functions that can find the optimal tax rate in one country, given the rate in the other.

However, when incorporating those functions into the model, the model will predict that Non-Havens will set their corporate tax rates unrealistically high, starting at an ETR of around 30% for a MR of 2.5% and only going up from there. For example, if the MR is 15%, it will say that the optimal ETR for the United Kingdom is 44.5%. Because the model only calculates optimal tax rates in response to profit shifting, it ignores exogenous factors that might disincentivize increasing tax rates. Therefore, the model is more realistic with assumption 4 in place than with the Kanbur-Keen response functions.

4.3. Data description

I use data collected by Tørsløv et al. (2018), written by authors from the Danish Ministry of Taxation and the University of California Berkeley. Tørsløv et al. (2018) is a comprehensive research paper that uses macrodata to map out the extent to which profit shifting is currently taking place in the world. The authors made a dataset containing their results freely available on https://missingprofits.world/. The data is organized by individual tax jurisdictions and contains estimates for how much profit is shifted to or from the jurisdiction, the average ETR for MNEs in that jurisdiction and the amount of pre-tax profits MNEs generated in that jurisdiction (for non-havens).

The methodology of Tørsløv et al. (2018) is as follows: The authors noticed that in tax havens, foreign firms have much higher ratios of pre-tax profits to wages than domestic firms and that in high tax jurisdictions, the opposite is true. This discrepancy is partly due to profit shifting since profit shifting is not tied to real economic activity and thus not inherently tied to a wage bill in the tax haven. By controlling for other factors that could affect the pre-tax profits to wage ratio, the authors could use that ratio to estimate the amount of profit shifted to each nation. Finally, the authors traced the shifted profits back to the jurisdictions where it was shifted from, using bilateral balance of payments data.

The research done by Tørsløv et al. (2018) is some of the most trusted work on profit shifting done in recent years. Despite being recently published, the paper has 520 citations on Google Scholar. For example, the OECD's impact evaluation (OECD 2020c) and Devereux et al. (2020) have used this dataset in their estimations of the effects of Pillar Two.

I add population statistics for each nation from the population dataset published by the World Bank (World Bank 2022). Since Tørsløv et al. (2018) use data from 2015, I used population data from each jurisdiction from the same year. Four tax havens – Anguilla, Bonaire, Jersey and Guernsey – do not have an entry in the World Bank database. It is important not to drop them since a substantial amount of profit gets shifted to these jurisdictions. Therefore, I added population statistics manually from local censuses (Anguilla Statistics Department 2015; Centraal Bureau voor de Statistiek 2022; Statistics Jersey 2022; States of Guernsey Data and Analysis Services 2015). Some of those statistics do not come from 2015, meaning that there is a minor discrepancy between their population in the data and their actual population in 2015. However, since these jurisdictions all have a very small population compared to the rest of the dataset, it is unlikely that this has any consequences.

Finally, I removed Turkey from the dataset. According to the Tørsløv et al. (2018) dataset, Turkey is not a tax haven, yet it has a lower ETR than the average ETR paid on shifted profits in tax havens. This fundamentally breaks the Kanbur-Keen model because $\tau_{NH} - \tau_H$ requires τ_H to be lower than τ_{NH}. There is no solution to this besides removing Turkey or setting profits shifted from Turkey to zero in all scenarios, which is effectively the same as removing it.

4.4. Using aggregates

In the Kanbur-Keen model, there are only two tax jurisdictions. In the real world, there are many more. This is a problem because $(\tau_{NH} - \tau_H)$ cannot be found on a simple nation-by-nation basis, because one jurisdiction's ETR needs to be compared to the ETR of the other jurisdiction.

The solution to this is to run the model on a country-by-country basis, but use a rate found through aggregation to replace the tax rate of the other jurisdiction. That means for every non-haven, the amount of profit shifted away after the introduction of a MR is calculated similarly to the basic model, except that the now non-existent τ_H is replaced with τ_{H_avg}, which is the average ETR in havens, weighted by the amount of profit shifted to each haven. This means that the average rates represent the global average percentage of tax paid on shifted profits. Likewise for each haven, τ_{NH} is replaced τ_{NH_avg}, calculated analogously. Before the reform, τ_{H_avg} is 5.8% and τ_{NH_avg} is 19.5%, but the model recalculates those values for every MR, to factor in that jurisdictions with an ETR below the MR are setting their ETR equal to the MR.

The drawback of this method is that the model is less sensitive to specifics of individual flows of profit shifting. For example, MNEs frequently shift profits generated in the European Union to tax havens inside the Union. Havens inside the EU typically have higher tax rates than Caribbean tax havens, but extensive economic openness within the EU makes shifting profits within the EU more accessible. So if MNEs in Germany, for example, primarily use The Netherlands and Luxembourg as their tax haven, they are facing a higher τ_H than the model suggests. This can be improved upon by computing all profit shifting flows between individual non-havens and individual havens, by using foreign affiliates statistics similar to what Tørsløv et al. (2018) did, and calculating changes in profit shifting per profit shifting flow. Doing this in addition to constructing the rest of the model is unfortunately beyond what one student can do in a few weeks.

4.5. Calculating post-reform amounts of profit shifting

In chapter 3.4. I have explained that using semi-elasticities is the preferable way to calculate changes in profit shifting, and this is displayed in the simple model, using a tax semi-elasticity found in the literature. However, tax semi-elasticity needs to be calculated endogenously in the complex model, because it uses real-world data. The reason for that is simple. There might be a difference between semi-elasticities that follow mathematically from the data and the 'consensus rate' of 0.8. In instances, the model will attribute the resulting difference in profit shifting to the introduction of the MR, thus creating biased results.

Semi-elasticities can be computed endogenously with $\varepsilon = \frac{S}{\Pi(\tau_{NH} - \tau_H)}$ [12] as the initial values of the right-hand side variables are known. Assuming all jurisdictions comply with the reforms, ε is constant across tax rate differentials. Changing amounts of profit shifted can be found with the function $S = \Pi * \varepsilon * (\tau_{NH} - \tau_H)$, where Π and ε are constant, where $(\tau_{NH} - \tau_H)$ is found with the tax rate differential calculation from 4.4 and where ε can be found using the initial amount of profit shifted. The new function takes the form of:

$$S_{i,t=1} = \frac{S_{i,t=0}}{\tau d_{t=0}} * \tau d_{t=1} \tag{9}$$

Where, *i* is the country in question, τd is the tax rate differential, calculated as described in 4.4. and *t* is the time period. *t=0* means prior to the introduction of a minimum tax rate, *t=1* means after the introduction of a minimum tax rate. Proof of equation 9 is in appendix 1.2. Π can be assumed as constant, even though annual profits are not constant in reality, because the model exclusively models profit shifting, meaning that factors that would change Π in reality, such as growth or decline of the economy or stochastic shocks, are not considered. In fact, the model is more accurate when holding Π constant as doing otherwise would mean a change in *S* as a consequence of a change in Π would be attributed to the introduction of the MR.

4.6. Model Limitations

This model is limited in several ways. Chapters 4.2. and 4.4. already deal with some limitations, but there are a few more that this chapter will discuss. Firstly, calculating the change in profit shifting through semi-elasticities is an imperfect method. In reality, MNEs factor in many variables in their decision to shift profits. Those include tax differentials but also concealment costs and stability or trustworthiness of the government and the financial system in the tax haven. Since the model uses real-world data, these other variables are, in a sense, factored in. For example, consider two tax havens, A and B. Both have the same ETR, but most MNEs prefer to shift profits to A over B, because they consider the financial system in A more stable which makes them more willing maintain an entity in A. The model will predict that even after Pillar Two is implemented, more profits get shifted to A than to B, because the model uses semi-elasticities and thus calculates new amount of shifted profit as a proportion of the original amount. However, because the decision to shift profits is based on a complex calculus of different factors, the real relationship between profit shifting and tax differential isn't necessarily one of gradual linear decline.

Notably, this is a problem for model outputs for very high minimum rates, such as 22.5% or 25%. Because there are some jurisdictions in the dataset with an ETR higher than 25%, the model will predict that MNEs will still attempt to shift some profits away from those jurisdictions even if the MR is set to 25%. Because the MR is 25% and because governments are assumed to not set tax rates below the MR, a quarter of all shifted profits end up as tax revenue for havens in this scenario. This revenue gain can be substantial if measured in terms of revenue gain per capita for tax havens with a relatively small population. In reality, it seems implausible that that this revenue gain will actually happen, because for MNEs, the cost of maintaining a profit shifting operation will likely outweigh the benefits of paying 25% tax instead of, say, 27% on a small fraction of their profits. This problem reduces the reliability of the model, but the extent of that problem is limited as it primarily seems to occur at a minimum rates that are much higher than the model would recommend.

Secondly, my model only models the revenue effects of Pillar Two. This specification is valuable and practical, for reasons explained in the previous chapters. But this is a double-edged sword. Because in

---

[12] $S = \Pi * \varepsilon * (\tau_{NH} - \tau_H)$ can be rewritten as $\varepsilon = \frac{S}{\Pi(\tau_{NH} - \tau_H)}$

reality, Pillar Two will not be passed in isolation. If it gets implemented, it will be in conjunction with Pillar One. Additionally, during the negotiation process, it is possible that proponents of Pillar Two offer certain concessions to jurisdictions that are least likely to profit from Pillar Two in order to get them to agree. Finally, Pillar Two will change investment decisions of MNEs which is not included in the model. The problem here is that even if the model is able to perfectly balance gains to havens and non-havens to make the trade-off described in Chapter 2.1., in reality, factors external to the model could affect that balance.

Thirdly, input data is based on estimates and is thus contestable. While the database from Tørsløv et al. (2018) is commonly used and based on thorough research, it is important to note that the values they find are estimates too. In particular, estimates for ETR vary widely in the literature. To illustrate, table 1 compares the ETR estimates for some jurisdictions from the database with estimates for ETR collected by Garcia-Bernando and Janský (2022). A discrepancy between ETR estimates in the model's input data and actual ETR of jurisdictions can hurt the predictive power of the model.

Jurisdiction	ETR from Tørsløv et al. (2018)	ETR from Garcia-Bernando and Janský (2022)
Netherlands	10.5%	6.8%
Switzerland	8%	5.5%
Japan	26.2%	20.5%
Brazil	19.5%	25.5%

Table 1. ETR estimates vary widely between datasets.

Finally, tax reform has a political dimension as well. Voters and politicians have opinions on tax reform based on other things than economic efficiency or tax revenue gain, for example ideology or trust in the international community. When the model predicts if a jurisdiction stands to gain or lose from Pillar Two, it does so exclusively from a tax revenue perspective. It cannot predict if the political leadership from that jurisdiction is willing to agree to Pillar Two. Therefore, the fact that the model can find Pareto improving rates does not mean that the model can predict whether or not Pillar Two will be implemented. It can only predict how large the tax revenue benefits will be and how they will be divided among jurisdictions.

5. Results

5.1. Aggregate tax revenue changes

Figure 2 shows the development of aggregate tax revenues of havens and non-havens, in response to minimum rates. The figure shows striking similarities with the output from the simplified model. A notable difference is that tax revenue increases immediately for any MR above 0. This is because the pre-reform ETR in Haven from the simplified model was 5%, but in reality, some notable tax havens like Bermuda or the British Virgin Islands actually have an ETR of 0%, meaning that the effects of a MR are noticeable even at very low rates. Another difference is that tax revenue in non-havens seems to increase exponentially, rather than linearly. This has as similar explanation; because many non-havens have an ETR below 25%, a high MR means more taxes paid on domestically declared profits as well as less shifted profit. Finally, the tax revenue of tax havens looks very modest. Note that the model only includes tax revenue from taxes paid on shifted profits. Because only about 5% of total corporate profits are shifted before the MR and because tax havens have lower tax rates than non-havens, it stands to reason that the total amount of taxes paid on shifted profits is very small compared to the total amount of corporate tax paid.

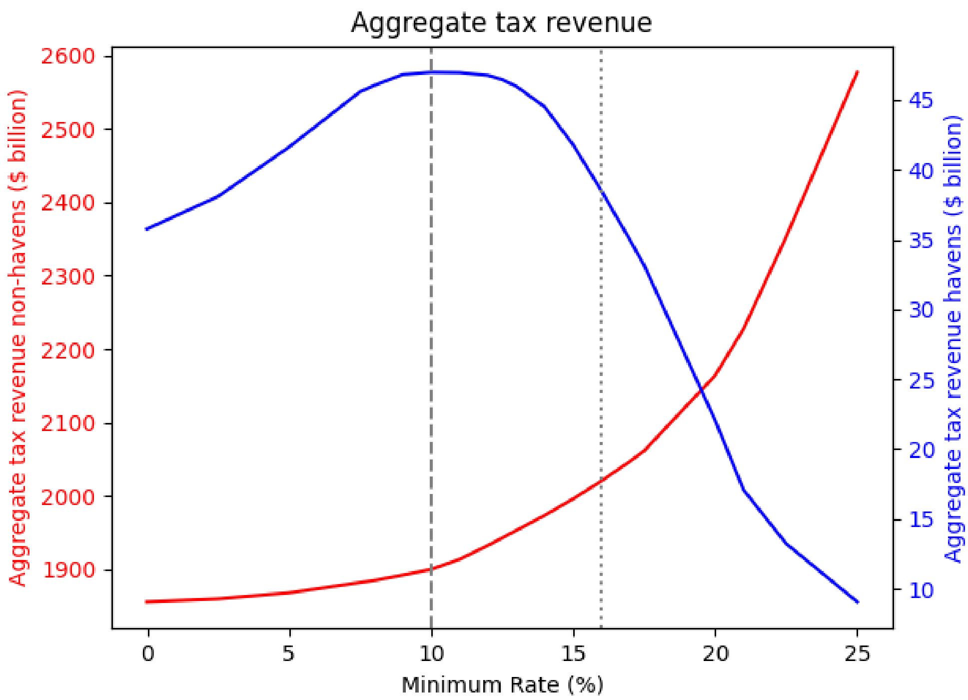

Figure 2: aggregate statistics.

If aggregates are assumed to represent all havens and non-havens, the Pareto efficient and maximum Pareto dominant MRs are 10% and 16% respectively. These are again represented by the dashed and dotted lines. Per capita figures lead to very similar recommendations. Figure 3 shows welfare gain per capita, defined in tax revenue gain as per the model assumptions, compared to the situation with no minimum tax. The intersection of both graphs, represented by the dashed-dotted line, is the point where per capita welfare gain in havens is equal to per capita welfare gain in non-havens. This is at approximately 16.25%, meaning that the maximum Pareto dominant rate provides the most equal division of benefits between havens and non-havens.[13]

The estimates from aggregate data are encouraging but present only a limited view of what is happening, as they are insensitive to differences between different jurisdictions within both groups. To illustrate, Figure 4 shows per capita revenue gain, identical to figure 3, but for a select group of individual jurisdictions. All non-havens gain between 0 and 250 dollar per person, increasing quasi-linearly with the height of the minimum rate. While there are noticeable differences between their marginal revenue gains from an increase in the MR, the differences between havens are even more striking. Guernsey benefits substantially more than any of the non-havens, even at an MR of 20%, while Singapore hardly benefits at all, even from the Pareto efficient rate, and very quickly stands to lose a lot from the reform.

[13] The reason why the most equal division of benefits is so close to the optimal rate for non-havens, rather than in the middle between the Pareto efficient and maximum Pareto dominant rate, is that per-capita revenue gain in non-havens is marginal compared to benefits for havens within their preferred range of minimum rates

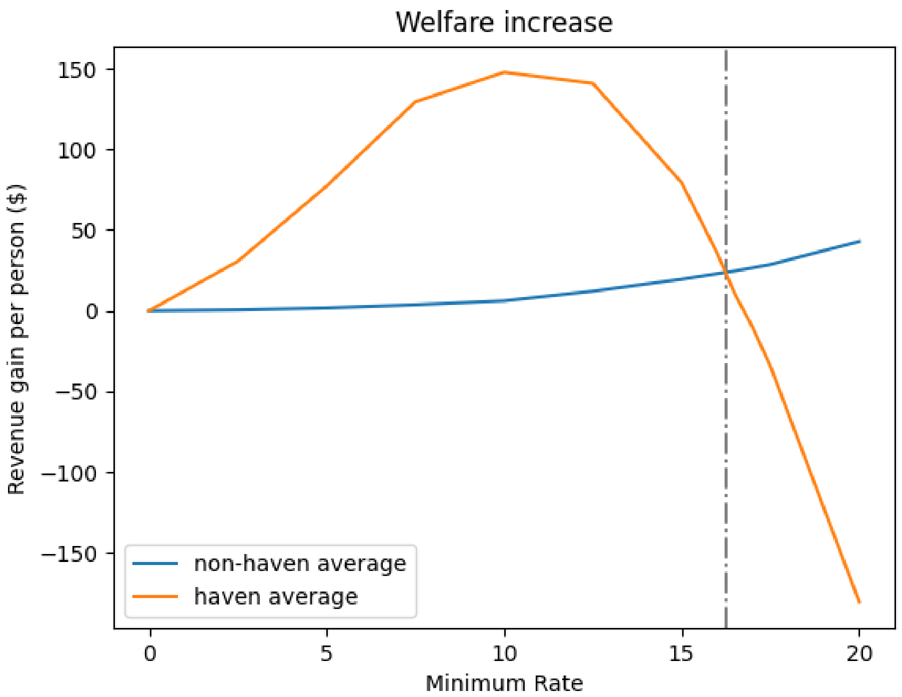

Figure 3. Welfare increase as revenue gain per capita

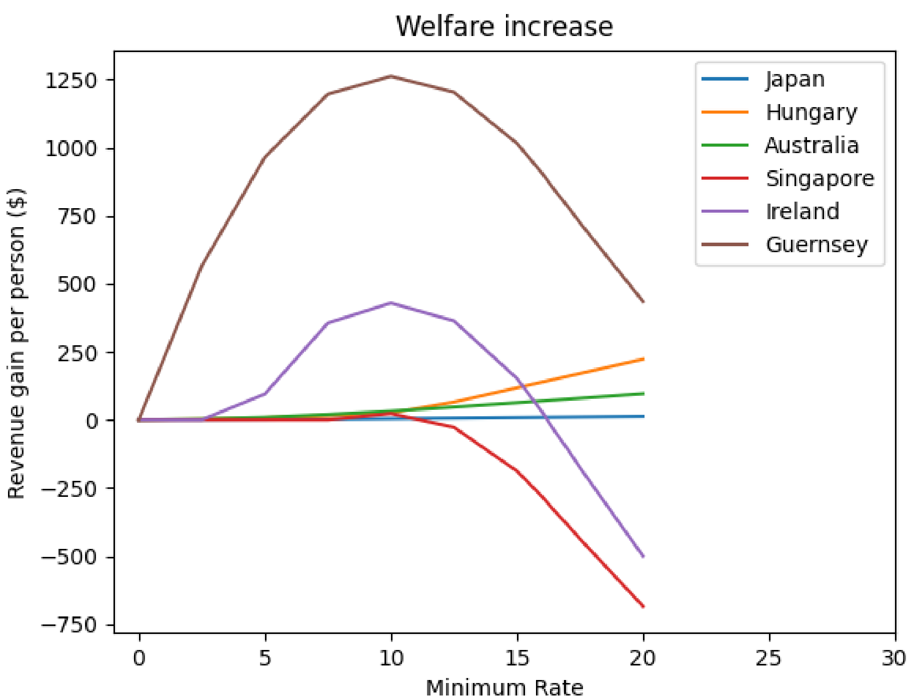

Figure 4. Effects on some individual jurisdictions

5.2. Breaking down havens

I now depart from aggregate statistics, o account for the differences between havens. Havens are the most important group of jurisdictions to analyse in-depth, because under the assumption that real economic activity is unchanged, non-havens stand to benefit regardless of how high the MR is set. The best way to represent incentives of individual havens is to show how many of them benefit from a given MR, since the incentive to not comply occurs when a jurisdiction is expected to lose from the reform. Figure 5 is constructed with that aim. For each level of minimum rate, it gives the percentage of havens that break even or improve compared to the world without a MR. At 10% or below, that is 100%, meaning that the real maximum Pareto dominant rate is 10%. At 11%, the Netherlands is the only haven that stands to lose from the reform. Switzerland and Singapore quickly follow suit.

The most salient result in figure 5 is the sudden drop at 15%. At 15%, only The Netherlands, Singapore, Guernsey and Switzerland lose from the reform. At 16%, 19 of the 39 tax havens stand to lose from the reform. This is important because a rate might still be acceptable if only a small number of havens stand to lose. For instance, they might still be persuaded to support and enforce the reform because they benefit from it in ways not captured by the model. Additionally, the enforcement problem from Hindriks and Nishimura (2022) is much less significant if cooperation only gets broken down in a small minority of cases.

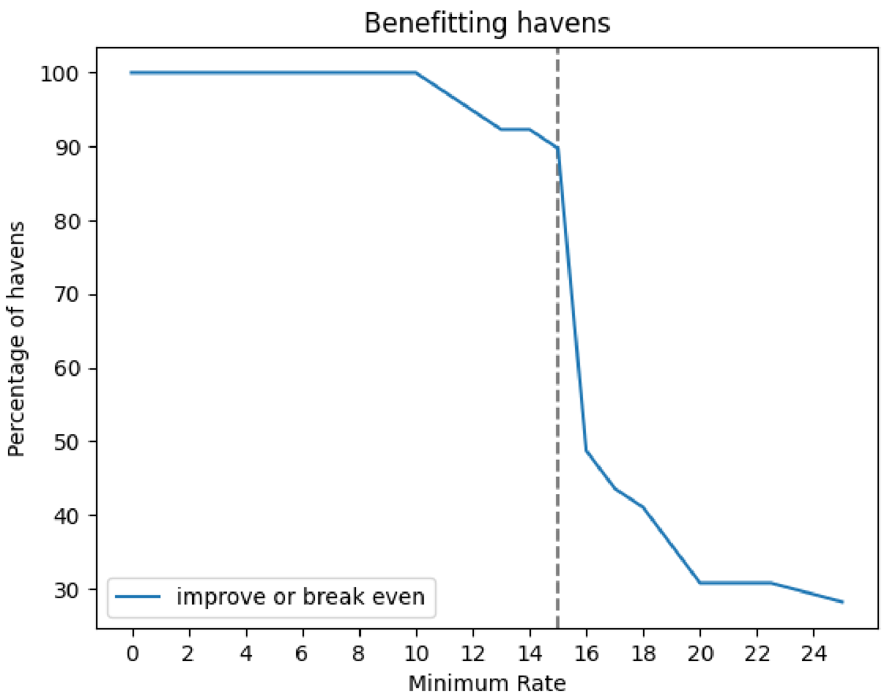

Figure 5: Percentage of benefitting havens

Figure 6 is a slight modification of figure 5. It adds the percentage of havens that stand to lose at least $100 per capita and the percentage that stand to gain at least $1000 per capita. This shows the difficulty of setting the minimum rate. At lower rates, a large minority of havens stands to gain $1000 per capita, which is substantially more than any non-haven, thus creating the situation Johannessen (2022) warns for where benefits disproportionally accrue in havens. However, if rates get set higher to avoid that, several havens stand to lose so much that they have a strong incentive to oppose the system, creating the situation Hindriks and Nishimura (2022) warn for. In figure 6, note that the numbers for per capita revenue gain of more than $1000 might be biased upwards for the higher range of minimum rates, as explained in chapter 4.6.

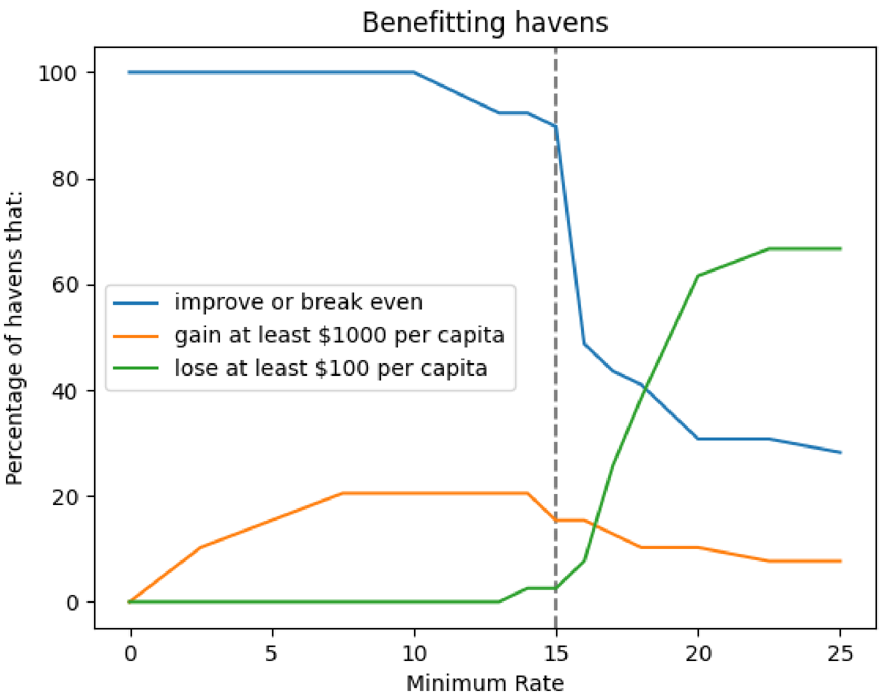

Figure 6 Expanded version of figure 5

The reason why this trade-off is so difficult to make is illustrated in figures 7 and 8. One group of havens, particularly havens with a relatively high population and ETR compared to other havens, hardly stands to gain anything,[14] while another group of havens, particularly small Caribbean havens with very low ETRs before the reform, stand to gain an unreasonable amount. These numbers are so high for two reasons. One, these jurisdictions have a relatively small population and because shifted profits are not real economic activity, the size of the havens' economy does not factor into how many profits can be shifted. Two, these jurisdictions have zero or near zero initial ETR, which resulted in profit shifting to a degree disproportional to their real economy, but it also means that their initial tax revenue was very small, making the increase induced by charging the minimum rate larger.

[14] Note that this is not a consequence of using aggregate tax rates, as data for each haven doesn't use aggregate rates for havens, only aggregate rates for non-havens. The mathematical reason behind this is that the initial tax differential is relatively low for havens with a relatively high ETR. Consequently, the model will predict a relatively large decrease in proportion of profit shifted for a marginal decrease in tax differential.

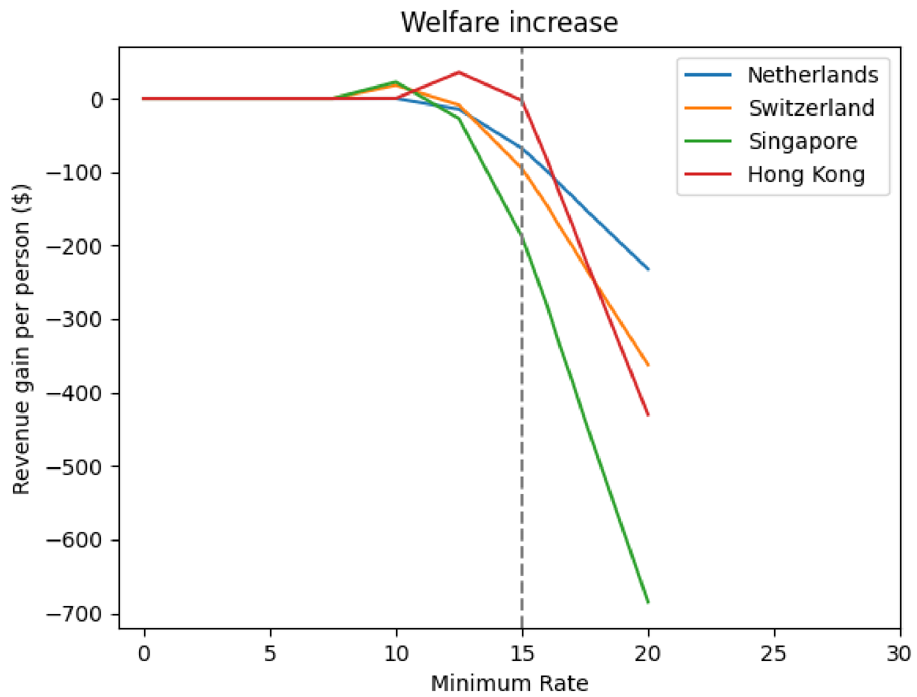

Figure 7. Havens at risk of losing too much

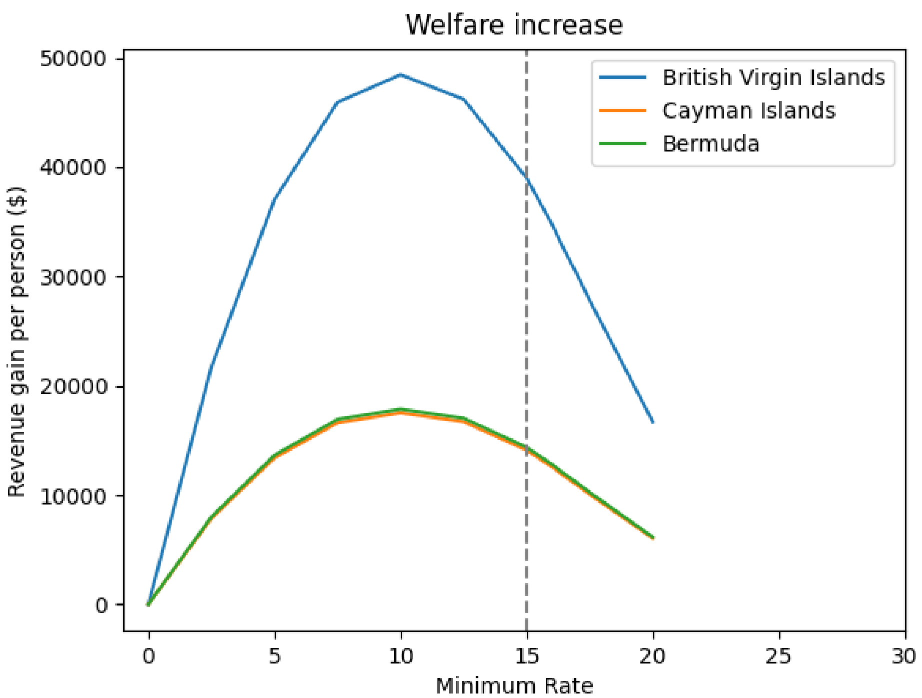

Figure 8. Havens at risk of gaining too much. Data for the Cayman Islands and Bermuda is very similar, hence why their curves are nearly identical.

6. Conclusion and discussion

6.1. Conclusion

Based on these results, I can be cautiously optimistic about the proposed 15% minimum rate. Based on aggregate statistics, the range of advisable minimum rates is 10%-16%, accounting for differences between tax havens shows that the real Pareto rates are much lower, with the maximum Pareto dominant rate being 10%. However, for 15% or lower, only a small number of jurisdictions will lose out from the reform, meaning that the political loss from choosing a rate that isn't Pareto improving compared to the world without a minimum rate is relatively small. Therefore, I conclude that the rate chosen by the jurisdictions participating in the Inclusive Framework seems to be a sensible one, and plausibly the most optimal rate. Especially since 15% is also very close to the optimal rate based on aggregate statistics and close to the most equal division of welfare gains between havens and non-havens. It is likely better than the maximum Pareto dominant rate because the model predicts a serious risk of benefits accruing disproportionally in havens, thus leaving the system at risk to the concern raised by Johannesen (2022). Because the maximum Pareto dominant rate of 10% is on average the optimal rate for havens, this risk is particularly high at that rate and a higher rate is likely preferred in practice.

Additionally, the model finds heterogenous revenue effects in havens, where havens with a relatively large real economy, such as Switzerland, Singapore and the Netherlands benefit only marginally from minimum taxation, but smaller tax havens, particularly in the Caribbean, stand to gain substantially. This is potentially encouraging since their policymakers will likely weigh incentives related to their real economy over the incentive to keep profit shifting. Regardless, this has two important implications. One, it narrows down the number of havens which are at risk of not complying with the proposal and which might need to be offered a benefit to offset this potential loss. Two, the Inclusive Framework should consider a way to prevent extreme revenue gain in tax havens from shifted profits to prevent the concerns from Johannesen (2022).

6.2. Discussion

While the results of this study support a minimum rate of 15%, they are by themselves not enough to provide a holistic policy recommendation on how high the minimum rate should be. That is because they are based on one dimension of impact and a narrow range of behavioural assumptions. Further research to explore other dimensions of this question is needed to come to a definitive answer to the question.

That is firstly in terms of further specifying the model. Firstly, this model assumes that all havens will choose to respond to the introduction of Pillar Two by setting their ETR equal to the minimum rate. In practice, that might not end up happening. For example, because of limited manpower in tax administration. The jurisdictions with an ETR of zero in the dataset are all very tiny Caribbean islands. They might lack the resources to set up a tax system that is capable of ensuring that all corporate entities in their jurisdiction pay a near-optimal amount of tax. As a consequence of this assumption not being met, a larger proportion of the tax revenue gains will materialize in non-havens than the model predicts. That is because tax revenue on low tax profits will now go to non-havens as a consequence of the income inclusion rule, rather than to havens as a consequence of them increasing their ETR. Secondly, the model suggests a large heterogeneity of havens. This suggests that making the model more capable of measuring effects on individual jurisdictions is a prudent next step. For example, by feeding the model with data that shows the source and destination country for each stream of profit shifting. Further improving the model like this will improve its strength.

Additionally, holistic policy advice needs to factor in the real economic dimension, for two reasons. One, it makes the effect of a minimum rate on non-havens more ambiguous, as a higher tax burden for MNEs adds an extra constraint on their real economic activity, which hurts the jurisdiction in which that economic activity takes place. Two, higher tax rates in havens not only reduces profit shifting to havens, it also reduces real economic activity in havens. Remember that Pillar Two is estimated to reduce FDI in tax havens by 4.0 to 7.3% (UNCTAD 2022). Because this is a welfare loss in havens that is not captured in my model, the real Pareto efficient and maximum Pareto dominant rates might be lower than my model predicts.

Finally, additional research is needed to model the effects of non-compliance. My model shows great heterogeneity between havens and consequently, it is unlikely that there is one level of the Minimum Rate that does not hurt at least one of them, other than rates that are low enough to trigger the problem described in Johannesen (2022). Therefore, a model of the consequences of non-compliance is needed to determine the optimal minimum rate.

7. References

Anguilla Statistics Department (2015). *The Fabric of our Households – Anguilla Population and Housing Census 2011*. https://unstats.un.org/unsd/demographic/sources/census/wphc/anguilla/AIA-2015-05-22.pdf

Brokelind, C. (2021). An Overview of Legal Issues Arising from the Implementation in the European Union of the OECD's Pillar One and Pillar Two Blueprint. *Bulletin for International Taxation*, *75*(5), 212-219.

Centraal Bureau voor de Statistiek (2022). *Caribisch Nederland; bevolking; geslacht, leeftijd, burgerlijke staat*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83698NED/table?ts=1549199994001

da Silva, B. (2020). Taxing Digital Economy: A Critical View around the GloBE (Pillar Two). *Frontiers L. China*, *15*, 111.

Devereux, M. P., Bares, F., Clifford, S., Freedman, J., Güçeri, İ., McCarthy, M., ... & Vella, J. (2020). The OECD Global Anti-Base Erosion Proposal. *PwC report, Oxford University Centre for Business Taxation, January*.

Finér, L., & Ylönen, M. (2017). Tax-driven wealth chains: A multiple case study of tax avoidance in the Finnish mining sector. *Critical Perspectives on Accounting*, *48*, 53-81.

Garcia-Bernardo, J., & Janský, P. (2022). Profit Shifting of Multinational Corporations Worldwide. ICTD Working Paper, 119, 1–72.

Hebous, S., Keen, M. (2022) : Pareto-Improving Minimum Corporate Taxation, CESifo Working Paper, No. 9633, Center for Economic Studies and ifo Institute (CESifo), Munich.

Heckemeyer, J. H., & Overesch, M. (2017). Multinationals' profit response to tax differentials: Effect size and shifting channels. *Canadian Journal of Economics/Revue canadienne d'économique*, *50*(4), 965-994.

Hindriks, J., & Nishimura, Y. (2022). *The Compliance Dilemma of the Global Minimum Tax* (No. UCL-Université Catholique de Louvain).

Hodgson, G. M. (2012). On the limits of rational choice theory. *Economic Thought*, 1(1, 2012).

Huizinga, H., & Laeven, L. (2008). International profit shifting within multinationals: A multi-country perspective. *Journal of Public Economics*, *92*(5-6), 1164-1182.

Johannesen, N. (2022) : The Global Minimum Tax, CESifo Working Paper, No. 9527, Center for Economic Studies and ifo Institute (CESifo), Munich.

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

Keen, M., & Konrad, K. A. (2013). The theory of international tax competition and coordination. *Handbook of public economics*, *5*, 257-328.

Massant, M., & De Broe, L. (2021). Are the OECD/G20 Pillar Two GloBE-Rules Compliant with the Fundamental Freedoms?. *EC Tax Review*, *30*(3).

Mooij, R. A. D., & Ederveen, S. (2008). Corporate tax elasticities: a reader's guide to empirical findings. *Oxford review of economic policy*, *24*(4), 680-697.

OECD (2020a), *Tax Challenges Arising from Digitalisation – Report on Pillar One Blueprint: Inclusive Framework on BEPS*, OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris, https://doi.org/10.1787/beba0634-en.

OECD (2021). *Tax Challenges Arising from the Digitalisation of the Economy – Global Anti-Base Erosion Model Rules (Pillar Two): Inclusive Framework on BEPS*, OECD, Paris, https://www.oecd.org/tax/beps/tax-challenges-arising-from-the-digitalisation-of-theeconomy-global-anti-base-erosion-model-rules-pillar-two.htm.

OECD, (2017). *Background Brief Inclusive Framework on BEPS*, Organisation for Economic Co-operation and Development, Paris, https://www.oecd.org/tax/beps/background-brief-inclusive-framework-on-beps.pdf.

OECD, (2020b). *Tax Challenges Arising from Digitalisation – Report on Pillar Two Blueprint: Inclusive Framework on BEPS, OECD/G20 Base Erosion and Profit Shifting Project*, OECD Publishing, Paris, https://doi.org/10.1787/abb4c3d1-en.

OECD, (2020c). *Tax Challenges Arising from Digitalisation – Economic Impact Assessment: Inclusive Framework on BEPS*, OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris, https://doi.org/10.1787/0e3cc2d4-en.

States of Guernsey Data and Analysis Services (2015). *Guernsey Quarterly Population, Employment and Earnings Bulletin.* https://gov.gg/article/151289/Annual-Electronic-Census-Report-and-Quarterly-Bulletins-2015

Statistics Jersey (2022). *Census 2021 First Bulletin*. https://www.gov.je/News/2022/Pages/FirstCensusResultsPublished.aspx

Tørsløv, T. R., Wier, L. S., & Zucman, G. (2018). The missing profits of nations (No. w24701). National Bureau of Economic Research.

UNCTAD (2022). *World Investment Report 2022: International tax reforms and sustainable investment,* UNCTAD – United Nations Publication. https://unctad.org/system/files/official-document/wir2022_en.pdf.

Wilson, J. D. (1986). A theory of interregional tax competition. *Journal of urban Economics*, *19*(3), 296-315.

World Bank, World Development Indicators (2022). *Population, total*. https://datacatalog.worldbank.org/search/dataset/0037712/World-Development-Indicators

Zodrow, G. R., & Mieszkowski, P. (1986). Pigou, Tiebout, property taxation, and the underprovision of local public goods. In *TAXATION IN THEORY AND PRACTICE: Selected Essays of George R. Zodrow* (pp. 525-542).

Appendix 1: mathematical proof

Appendix 1.1

Proof that $\varepsilon = \frac{1}{\delta}$, *for* S = S*

ε is the percentage increase of profit shifting for each percentage increase of the tax rate differential, therefore:

$$S = \Pi * \ \varepsilon * (\tau_{High} - \tau_{Low})$$

Equation 4 is:

$$S^{*} = \mathit{\Pi} * \frac{\tau_{High} - \tau_{low}}{\delta}$$

The text says equivalent relationship of ε to S as δ to S^*, therefore S = S^*, therefore:

$$\Pi * \ \varepsilon * \ (\tau_{High} - \tau_{Low}) = \mathit{\Pi} * \frac{\tau_{High} - \tau_{low}}{\delta}$$

Dividing both sides by $\Pi(\tau_{High} - \tau_{Low})$ yields:

$$\varepsilon = \frac{1}{\delta}$$

Appendix 1.2.

Proof that $S_{i,t=1} = \frac{S_{i,t=0}}{\tau d_{t=0}} * \tau d_{t=1}$ follows from $S = \Pi * \ \varepsilon * (\tau_{High} - \tau_{Low})$, if $\mathit{\Pi}_{i,\ t=0} = \mathit{\Pi}_{i,\ t=1}$ and $\varepsilon_{i,\ t=0} = \varepsilon_{i,\ t=1}$:

$$S_{i,\ t=0} = \Pi_i * \varepsilon_i * d\tau_{i,\ t=0} \xrightarrow{rewrite\ to} \varepsilon_i = \frac{S_{i,\ t=0}}{\Pi_i d\tau_{i,\ t=0}}$$

$$S_{i,\ t=1} = \Pi_i * \varepsilon_i * d\tau_{i,\ t=1} \xrightarrow{Substitute\ for\ \varepsilon_i} S_{i,\ t=1} = \Pi_i * \frac{S_{i,\ t=0}}{\Pi_i d\tau_{i,\ t=0}} * d\tau_{i,\ t=1}$$

$$S_{i,\ t=1} = \frac{\Pi_i S_{i,\ t=0}}{\Pi_i d\tau_{i,\ t=0}} * d\tau_{i,\ t=1} \xrightarrow{yields} S_{i,\ t=1} = \frac{S_{i,\ t=0}}{d\tau_{i,\ t=0}} * d\tau_{i,\ t=1}$$

Appendix 2: Reproduction Guide

Attached to this thesis is a zip file containing the datasets of my models' inputs and outputs and the Python code used to create those. The dataset called 'Thesis Harmen inputs' contains the model inputs. This data is created by running the code in 'Thesis Harmen Data Merger'. The source of the data and the data editing process are discussed in simple terms in chapter 4.3. 'Thesis Harmen outputs' contains the model outputs. The column names represent the minimum rate for which the numbers below apply. For example, in 'Revenue Havens', the column marked 10 displays tax revenue on shifted profits in tax havens under a minimum rate of 10%. All figures in this dataset are in billions of dollars.

The Python code is saved in .txt files, but it can be copied into any application capable of running Python code. I personally used PyCharm Community Edition, which I can recommend. Note that in order to run the Python code, you need change the file directory to the location where you saved the file. Running the data merger file requires you to download the data I used from Tørsløv et al. (2018) from missingprofits.world. The necessary files are 'Main data' and 'Replication Guide Figures'.

Addtionally, you need population data from the World Bank. A link to the full World Development Indicators dataset is in the References section of this thesis, however, I highly recommend that you only download population statistics since the dataset is rather large and this is the only part you need to run the file. The population statistics section of the dataset can be obtained from https://data.worldbank.org/indicator/SP.POP.TOTL.