European University Institute

EUI Working Papers

ECO 2009/15
DEPARTMENT OF ECONOMICS

Heterogeneous Firms, `Profit Shifting' FDI and International Tax Competition

Sebastian Krautheim and Tim Schmidt-Eisenlohr

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
DEPARTMENT OF ECONOMICS

Heterogeneous Firms, 'Profit Shifting' FDI and International Tax Competition

SEBASTIAN KRAUTHEIM
and
TIM SCHMIDT-EISENLOHR

EUI Working Paper **ECO** 2009/15

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

Heterogeneous Firms, 'Profit Shifting' FDI and International Tax Competition*

Sebastian Krautheim

Paris School of Economics and European University Institute

Tim Schmidt-Eisenlohr

European University Institute

February 27, 2009

Abstract

We develop a stylized model of international tax competition between a large country and a tax haven. In the large country, firms in a monopolistically competitive industry generate positive profits which can be taxed by the government. Firms have heterogeneous productivity levels and can choose to undertake 'profit shifting' FDI in order to benefit from lower tax rates abroad. We find that economies with a low degree of firm heterogeneity and a high degree of monopolistic market power are less affected by international tax competition. They face lower outflows of the tax base and can set higher tax rates.

JEL: *F23, H25, H87*

Keywords: *heterogeneous firms, monopolistic competition, tax competition, tax havens*

* We are grateful to Giancarlo Corsetti, Omar Licandro and Morten Ravn for constant advice. We would also like to thank Jonathan Eaton, Vincent Rebeyrol, Karolina Ekholm, guest researchers at the Deutsche Bundesbank, participants at the EUI Macro Group and the PSE International Trade lunchtime seminar for fruitful discussions and comments. This work was supported by the Région Ile-de-France.

1 Introduction

International financial integration enables multinational companies to use international tax planing strategies to reduce their tax payments. As a consequence, when choosing their tax policy, governments are competing for a potentially mobile tax base.

Very few studies have considered the role of firm heterogeneity in the analysis of international tax competition.[1] Recent empirical evidence shows, however, that firm heterogeneity does matter for the extent of international tax optimization. Using affiliate-level data of American firms, Desai, Foley, and Hines (2006) show that firm size, internationalization and R&D intensity increase the probability of using tax haven operations. [2]

This paper proposes a fully solvable model of international tax competition with heterogeneous firms and monopolistic competition. The model allows to analyze how the main policy tradeoffs in international tax competition are affected by firm heterogeneity and monopolistic market power. These two factors determine the distribution of profits (the tax base) across firms. This distribution in turn is key for optimal government policies as some firms are more prone to shifting profits abroad than others. The analysis reveals that the effects of international tax competition are strongest in economies with a low degree of firm heterogeneity (relatively many productive firms) and high substitutability across goods (low monopolistic market power).

The model features a large country, in which firms are active in a monopolistically competitive industry and earn positive profits. These profits can be taxed by the government. In order to avoid taxation at home, firms can become multinationals by opening an affiliate in a foreign jurisdiction, the 'tax haven'. This tax haven is small in the sense that it does not have an industry of its own producing differentiated products. Using methods of profit shifting, multinationals can transfer all profits to their affiliates and thus effectively pay taxes in the tax haven only.[3]

The governments of the large country and the tax haven set their tax rates non-cooperatively. In the unique Nash-equilibrium of the model, the tax haven sets a tax rate below the one set by the large country. This gives firms an incentive to do 'profit shifting' FDI. While the fixed cost of opening an affiliate in the tax haven is the same for all firms, the gains from profit shifting depend on the level of profits a firm is making. In line with the findings of Desai, Foley, and

[1] Some notable exceptions will be discussed below.

[2] They analyze data on American multinational firms from the Bureau of Economic Analysis annual survey of U.S. Direct Investment Abroad for the years 1982 to 1999. Grouping countries with US affiliates into tax havens and non-havens allows them to find correlations between tax haven activities and firm level characteristics.

[3] In order to keep the analysis focused, we only allow for this 'profit shifting' aspect of FDI. We only consider the case where firms can shift their total profits abroad. Introducing partial profit shifting would neither affect the mechanisms in the model nor the qualitative results.

Hines (2006), in equilibrium the most productive (and thus largest and most profitable) firms become multinationals while less productive firms continue to pay taxes at home.

The higher the tax rate of the large country relative to the tax rate of the tax haven, the more firms create an affiliate in the tax haven to pay taxes abroad. An increase in the tax rate in the large country thus has two opposing effects on tax revenues at home. On the one hand it increases revenues per unit of profits taxed (intensive margin), on the other hand it decreases the size of the tax base itself by making profit shifting more attractive (extensive margin).

A key determinant of the optimal tax policies is how strongly the tax base reacts to tax differences. Since the most productive firms earn the highest individual profits they are the first to do 'profit shifting' FDI. When these firms account for a large fraction of aggregate profits, the tax base reacts strongly to tax differences. Whether this is the case depends on industry- and market structure. For the industry structure, a high degree of firm heterogeneity implies a larger number of very productive firms. In addition, these high productivity firms have relatively high individual profits when the degree of substitutability between goods is high, i.e. monopolistic market power is low. Thus high firm heterogeneity and low monopolistic market power imply strong reactions of the tax base to tax differences.

In this case, the large country suffers substantial outflows of tax base. The tax haven gains as it can set a relatively high tax rate and still attract a considerable fraction of the tax base.

When instead there is low firm heterogeneity and high monopolistic market power, the tax base does not react strongly to tax differences and the large country is 'protected' from international tax competition. It can set a relatively high tax rate without loosing much of its tax base. The tax haven is forced to undercut the large country strongly in order to attract some of the tax base.

This paper relates to two different strands of the literature on international tax competition. One is the Public Finance literature on international tax competition which builds on the seminal contributions of Zodrow and Mieszkowski (1986) and Wilson (1986) (see e.g. Wilson (1999) or Fuest, Huber, and Mintz (2005) for surveys). This literature mainly uses models with identical firms under perfect competition. Usually there are two factors of production: labor and capital. The former is in general assumed to be immobile and the latter to some degree mobile between countries. A benevolent government sets a capital tax rate to maximize welfare of its citizens. It faces a trade-off between increasing its budget for public good provision and the capital outflow implied by a higher tax rate. In equilibrium governments set inefficiently low tax rates.

Bucovetsky (1991) extends the model to asymmetric countries. He finds that in the Nash equilibrium the smaller country sets a lower tax rate and thus has net capital inflows. When the size difference is sufficiently large the small country can be better off under tax competition than when the two countries merge, while the large country always looses from tax competition.
Our analysis has some similarities with Bucovetsky (1991) as we consider 'very' asymmetric countries, a large country and a tax haven. Thus, some of the mechanisms are the same. Also in line with this strand of literature we allow for one single tax instrument only. We do, however, not consider per unit capital taxes but a proportional profit tax, which we believe to be the relevant instrument to study.
Building on this literature, Burbidge, Cuff, and Leach (2006) introduce a type of firm heterogeneity into a model with perfect competition, immobile labor and mobile capital. They model firm heterogeneity as an ideosyncratic, exogenous comparative advantage in one of the locations. So firms are heterogeneous in the sense that they are more productive in one country or the other. This is reflected by a country-firm specific TFP term in a decreasing returns to scale production function. Their analysis shows that these location-specific productivities can lead firms owned by residents of the home country to locate in the foreign market. The type of heterogeneity they use is fundamentally different from Melitz (2003) where independent of the location of their production firms differ in productivity and thus sales and profits.
Our model is also related to the literature on tax competition in a 'New Economic Geography' (NEG) context (see e.g. Baldwin and Krugman (2004), Borck and Pflueger (2006) and Ottaviano and van Ypersele (2005)). The analysis of NEG models is usually focused on the effect of taxation and tax competition on the allocation of monopolistically competitive firms between countries of different sizes. Due to the well known 'home-market effect' (Krugman (1980)), the larger country attracts more firms and can thus afford to set higher taxes on capital than the smaller country. The larger country thereby taxes capital owned by agents in the foreign market.[4]

At first sight it might seem to be a straightforward generalization of NEG models of tax competition to introduce firm heterogeneity as in Melitz (2003). To the best of our knowledge, however, there have not been many attempts to do so in the literature. Baldwin and Okubo (2008) outline a basic NEG model with tax competition and heterogeneous firms. However, they do not derive the equilibrium of their model. Instead, they focus their analysis on the

[4] Sato and Thisse (2007) show that in a setting with imperfect matching between firms and workers, this home market effect can be reversed.

investigation of a trade-off between two different tax instruments for an exogenous (possibly off-equilibrium) tax difference.

Davies and Eckel (2007) also propose an NEG-type model of tax competition with heterogeneous firms. In order to achieve tractability they make a particular assumption on the ownership structure of firms and thereby on the effect of profits and taxation on demand. They assume that firms belong to individual entrepreneurs which do not have an individual demand of their own. Once entrepreneurs change the country (to maximize profits of their firms), they pay taxes in that country and since they do not consume, profits go to the consumers in that country. With this ownership structure it seems unclear why firms would maximize profits at all as their owners do not benefit from them. This might also make the model less useful for the analysis of the effect of profit taxes.

In our paper we use the more standard assumption that firms belong to consumers. To the best of our knowledge our paper provides the first model of international tax competition with heterogenous firms and monopolistic competition as in Melitz (2003) and standard ownership assumptions that is fully solvable in closed-form. To achieve tractability, we abstract from several issues that are crucial in the analysis of the NEG literature, but not in our context: we consider a case of 'very' asymmetric countries considering a large country and tax haven. We are not analyzing firm creation or endogenous (re)location of firms but focus on existing firms. We abstract from international trade and focus on 'profit shifting' FDI. Furthermore, we assume quasi-linear preferences that are linear in a homogeneous good and the public good. In this setup both the total expenditure on differentiated goods and the aggregate price level are independent of the tax rates. While these are strong simplifications, we believe them to be appropriate for the purpose of the paper. They allow us to solve the model in closed form and to investigate the role of firm heterogeneity and monopolistic market power in international tax competition further than has been done so far in the literature.

The remainder of the paper is structured as follows. Section two presents the case of a large country in financial autarky. In section three the tax haven and 'profit shifting' FDI are introduced. In section four the best response functions of the international tax game are derived. Section five analyzes the equilibrium. Section six discusses the main implications of firm heterogeneity for international tax competition. Section seven concludes.

2 Financial Autarky

We first outline the structure of the large country in financial autarky. It is endowed with L workers. There are two sectors, one producing varieties of a differentiated good and one producing a homogeneous good with constant returns to scale. The homogeneous good is used as the numeraire with its price normalized to one. As standard in such a setting only equilibria are considered in which the homogeneous good is produced. Wages are normalized to unity. Labor is supplied inelastically and is the only input in production. There is a fixed and exogenous measure of firms which are owned by consumers in the large country.

Preferences: The workers are all identical and share the same quasi-linear preferences over consumption of the two goods and a good provided by the government:

$$U = \alpha \; lnQ + \beta \; G + q_0 \qquad with \qquad Q = \left(\int_{\Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \tag{1}$$

Where $q(\omega)$ is the quantity consumed of variety ω. The elasticity of substitution between varieties is given by $\sigma > 1$ and Q thus represents consumption of a preference weighted basket of differentiated goods. G is the quantity of a public good provided by the government. The consumption of the numeraire good is given by q_0. α and β are parameters with $0 < \alpha < 1 < \beta$.[5] Taking prices and available varieties as given, the demand for one particular variety takes the usual form:

$$q(\omega) = \frac{p(\omega)^{-\sigma}}{P^{1-\sigma}} \; \alpha. \tag{2}$$

Where $p(\omega)$ is the price of variety ω, the aggregate price index of the differentiated goods sector is given by $P = \left(\int_0^{a_m} p(a)^{1-\sigma} dG(a) \right)^{\frac{1}{1-\sigma}}$ and $\alpha = P \; Q$ is the overall expenditure on goods in this sector.[6]

The government: The only tax instrument of the government is a proportional tax on the profits of firms in the home country.[7] Tax income can be used to provide government services G

[5] These preferences are similar to those used in Baldwin and Okubo (2006) and Baldwin and Okubo (2008). To generate demand for the public good we add the term $\beta \; G$. The public good thus enters utility in the same way as the numeraire good except that it yields higher per unit utility ($\beta > 1$).

[6] $G(a)$ is the distribution function of cost levels of firms and a_m is the maximum cost level.

[7] This is a stylized way of looking at corporate taxation based on the 'source principal'. By allowing for one tax instrument only, we place our model in the tradition of the large literature on tax competition surveyed by Wilson (1999). Following Hamada (1966), some authors add a lump sum tax on labor, which usually leads to

to the consumers. The government can transform one unit of the numeraire good into one unit of the government services. The government is assumed to maximize welfare of its own citizens.

Firms: In the homogeneous good sector firms produce with a constant returns to scale technology and earn zero profits.

Similar to Chaney (2008), there is a fixed and exogenous measure of firms in the differentiated good sector that is without loss of generality normalized to one. Each firm produces a different variety. Firms differ in their levels of marginal cost, which is constant for each firm. We assume that these marginal cost levels reach over an interval of $[0, a_m]$ and are following a Pareto distribution with the distribution function given by

$$G(a) = \left(\frac{a}{a_m}\right)^{\gamma}$$

with a_m defined as the highest marginal cost i.e. the lowest marginal productivity. The degree of firm heterogeneity (technically: the variance of the distribution) is determined by a_m and the shape parameter γ of the Pareto distribution.[8] As standard in the literature, we assume $\gamma > (\sigma - 1)$ in order to assure that aggregate profits are finite. There is no fixed cost of production for firms, so that in equilibrium all firms produce.[9]

Firms in the differentiated good sector charge a constant mark-up over marginal cost:

$$p(a) = \frac{\sigma}{\sigma - 1}\, a. \tag{3}$$

The level of the mark-up depends on the elasticity of substitution between varieties. When σ is high, firms have a low degree of monopolistic market power and can only afford to charge a low mark-up.

A firm's gross profits are given by $\pi(a) = r(a)/\sigma$ which implies

$$\pi(a) = a^{1-\sigma}\, T_1 \quad with \quad T_1 = \frac{\alpha}{\sigma}\left(\frac{\gamma - (\sigma - 1)}{\gamma}\right) a_m^{\sigma-1}. \tag{4}$$

negative tax rates on capital (see for example Burbidge, DePater, Myers, and Sengupta (1997) or Ottaviano and van Ypersele (2005)).

[8]The parameter γ affects both the mean and the variance of the cost distribution. When we analyze the role of firm heterogeneity for tax competition we isolate the effect γ has via the variance by considering a mean preserving spread. For the time being just note that for a given a_m a change in γ always implies a change in the degree of firm heterogeneity.

[9]We abstract from the fixed cost of production for notational simplicity. Since we do not allow for firm entry or exit the introduction of a fixed cost of production would not qualitatively change the results.

T_1 is a constant that only depends on parameters of the model and the price index, which is also constant.[10]

Net profits are given by $\pi(a)^{net} = (1-t)\ \pi(a)$, where $t \in [0,1]$ is a tax rate set by the government and taken to be exogenous by the firm. Firm choices that maximize gross profits also maximize net profits. The tax is thus not distorting the optimal behavior of the firms.[11]

The tax base - aggregate profits: In financial autarky all firms pay taxes at home. The tax base is thus given by aggregate profits of firms $\Pi_H^A = \int_0^{a_m} \pi(a)\ dG(a)$. Evaluating the integral using (4) leads to

$$\Pi_H^A = \frac{\epsilon+1}{\epsilon}\ a_m^{1-\sigma}\ T_1 = \frac{\alpha}{\sigma} \tag{5}$$

or, expressed differently for later reference, $\Pi_H^A = T_2\ T_1^{-\epsilon}$. For notational convenience we have defined

$$\epsilon \equiv \frac{\gamma}{\sigma-1} - 1 \qquad and \qquad T_2 \equiv \frac{\epsilon+1}{\epsilon}\ T_1^{\epsilon+1}.$$

It is of interest to note that ϵ combines two of the crucial parameters of the model: the shape parameter of the cost distribution and the elasticity of substitution between varieties. Recall that above, we have assumed that $\gamma > (\sigma-1)$ which implies $\epsilon > 0$.

Optimal tax rate in autarky: Households have income from labor and receive the net profits of firms in their country. In autarky the aggregate income I^A of consumers is thus

$$I^A = L + (1-t_H^A)\Pi_H^A.$$

Welfare in financial autarky is then given by:

$$U^A = \bar{U} + (1-t_H^A)\ \Pi_H^A + \beta\ t_H^A\ \Pi_H^A. \tag{6}$$

Where $\bar{U} \equiv \alpha\ ln\left(\frac{\alpha}{P}\right) - \alpha + L$ collects terms that are unaffected by the taxation decision. The first term in $\bar{U}$ reflects utility of consuming the basket of differentiated products, α reflects the

[10]We have $T_1 = \left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1} \frac{\alpha P^{\sigma-1}}{\sigma}$. The price index is defined as $P = \left(\int_0^{a_m} p(a)^{1-\sigma} dG(a)\right)^{\frac{1}{1-\sigma}}$. Evaluating the integral using (3) leads to $P = \frac{\sigma}{\sigma-1}\ a_m \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{1-\sigma}}$ which is a constant. This implies $T_1 = \frac{\alpha}{\sigma}\left(\frac{\gamma-(\sigma-1)}{\gamma}\right) a_m^{\sigma-1}$.

[11]If we allowed for firm entry and exit this statement would still be true for active firms, but the tax rate would affect the set of active firms.

opportunity cost of consuming this basket instead of consuming the homogeneous good. L is labor income which is spent either on the basket or on the homogeneous good. The last term in (6) represents the part of aggregate firm profits that is payed to the government and transformed into the government service. The profits retained by consumers are used for consumption of the numeraire good.

The following proposition states the main result of this section.

Proposition 1 *In financial autarky the welfare maximizing tax rate of the large country is given by* $t_H^A = 1$.

Proof: This follows directly from the fact that $\bar{U}$ is constant, $\Pi_H^A > 0$ and $\beta > 1$. **q.e.d.**

While this result is of course very stylized, it provides a good benchmark for the analysis of the tax haven below.

3 Introducing 'profit shifting' FDI

We now consider the case where the large country described above is financially integrated with a small independent jurisdiction: the tax haven. Financial integration in our context means that firms in the large country have the possibility to become multinationals by opening an affiliate in the tax haven and shifting their profits from the headquarters to the affiliate. These profits are then taxed according to the tax rate in the tax haven, but not at home, where the firm declares zero profits. Opening an affiliate in the tax haven requires paying a fixed cost f_t.

Individual firm behavior and the tax base: Whether an individual firm chooses to pay the fixed cost of shifting profits abroad depends on the tax differential and on the level of profits the firm generates. The lower the marginal cost of a firm, the higher are the firm's profits and thus the more likely it is that the firm chooses to pay the fixed cost of 'profit-shifting' FDI.

We define the 'profit shifting cutoff cost level' as the cost level a^* for which a firm is indifferent between paying taxes at home and paying taxes in the tax haven. This cost level is determined by the following condition:

$$(1 - t_H)\ \pi(a^*) = (1 - t_X)\ \pi(a^*) - f_t.$$

Where $\pi(a^*)$ are gross profits of a firm with marginal cost of a^*, t_H is the domestic tax rate and t_X is the rate set by the tax haven. This condition only holds with equality if the tax difference $\rho = t_H - t_X$ is positive. When the tax difference is zero or negative no profit shifting takes place which corresponds to the financial autarky case described above. Rewriting the tax evasion cutoff condition gives

$$\pi(a^*) = f_t/\rho \tag{7}$$

and the 'tax evasion cost cutoff level' is

$$a^* = \left(\frac{\rho\, T_1}{f_t}\right)^{\frac{1}{\sigma-1}}. \tag{8}$$

We can now state the following Proposition:

Proposition 2 *Under financial integration when firms have the possibility to do profit shifting FDI, the most productive firms (with a cost level below a^*) self-select into profit-shifting FDI.*
Proof: This follows directly from (8). **q.e.d.**

This Proposition states an important result of the model: the most productive firms self-select into profit shifting FDI. In the model with heterogeneous firms, the mass of firms is thus endogenously split into multinationals and domestic firms. Each firm has the possibility to shift profits abroad, but only the most productive firms actually decide to do so. This pattern is in line with the empirical evidence on the determinants of the use of tax haven operations (see Desai, Foley, and Hines (2006)).

Tax base and number of multinationals: The tax base in the home country is given by aggregate profits of firms that have not become multinationals and thus pay taxes at home: $\Pi_H = \int_{a^*}^{a_m} \pi(a)\ dG(a)$. The tax base taxed in the tax haven is given by $\Pi_X = \int_0^{a^*} \pi(a)\ dG(a)$. Evaluating the integrals leads to

$$\Pi_X = \rho^{\epsilon} f_t^{-\epsilon} T_2 a_m^{-\gamma} \tag{9}$$

$$\Pi_H = \Pi_H^A - \Pi_X = T_2\left(T_1^{-\epsilon} a_m^{1-\sigma} - \rho^{\epsilon} f_t^{-\epsilon} a_m^{-\gamma}\right) \tag{10}$$

The measure of firms paying taxes in the tax haven is given by the measure of firms with marginal cost level below the cutoff level.

$$N_x = G(a^*) = (a^*)^{\gamma}\, a_m^{-\gamma} \tag{11}$$

Household Income: In addition to their income from labor, households receive the net profits of firms paying taxes at home and of firms paying taxes in the tax haven. Under financial integration, the aggregate income I of consumers is thus given by

$$I = L + (1 - t_H)\Pi_H + (1 - t_X)\Pi_X - N_x f_t. \tag{12}$$

Where the last term accounts for the fact that net profits of firms paying taxes in the tax haven are also net of the fixed cost payed to become a multinational.[12]

Governments: Under financial integration governments have to take into account the tax rate set in the other legislation. Taxes are set in a simultaneous one-shot game.[13] To analyze the tax game, we first derive the best response functions of the two governments.

4 Welfare Maximization under Financial Integration

Just as in financial autarky, the only variable governments can set are the profit tax rates in their legislations. In this section we derive the best response functions of the two governments in the international tax game. We start with the tax haven and then turn to the large country.

4.1 Optimization of the Tax Haven

The structure of the tax haven is kept as simple as possible. It does not have a tax base of its own. Its only source of revenue stems from taxing multinational companies that have an affiliate in the tax haven. Taking the tax rate in the large country as given, the tax haven maximizes total revenue $V = t_X \Pi_X$.[14] The attracted tax base Π_X will only be positive if the tax haven sets a lower tax rate than the large country. Thus for any given (positive) tax rate of the large country $\bar{t}_H$, it will always be optimal for the tax haven to undercut, so that $\rho > 0$. Revenue maximization leads to

$$t_X = min\left\{\frac{\rho}{\epsilon}; \frac{1}{\epsilon + 1}\right\} = \frac{\bar{t}_H}{\epsilon + 1} = \frac{\sigma - 1}{\gamma}\, \bar{t}_H. \tag{13}$$

[12] We implicitly assume, that the fixed cost of tax avoidance is not tax-deductible, i.e. profits are not taxed net of these fixed costs.

[13] The case where one country has a first mover advantage is discussed in the appendix.

[14] We can think of the tax haven as the limit case of a very small country. The measure of differentiated goods firms is proportional to the mass of consumers, which are close to zero. In this case the tax haven's own tax base is 'almost zero'. The same holds for demand for the differentiated good (imported from the large country with zero trade cost). In this case welfare maximization of the government in the tax haven boils down to maximization of tax revenues.

Note that $\epsilon + 1 = \gamma/(\sigma - 1) > 1$. The tax haven sets a tax rate that is a constant fraction of the rate of the large country. The extent to which the tax haven undercuts the large country is determined by the shape parameter of the cost distribution and the elasticity of substitution. We can now state the following proposition:

Proposition 3 *Under financial integration when firms have the possibility to do profit shifting FDI,*

(i) the best response of the tax haven for a given $\bar{t}_H > 0$ *is given by* (13) *and implies that the tax haven always undercuts the large country.*

(ii) the undercutting is the stronger the higher the shape parameter γ *and the stronger the market power of individual firms (lower* σ*).*

(iii) the best response of the tax haven for $\bar{t}_H = 0$ *is indeterminate.*

Proof: *(i)* follows from the fact that for $\bar{t}_H > 0$, a tax rate of $t_X \geq \bar{t}_H$ implies $\Pi_X = 0$ and thus $V = 0$, while any $0 < t_X < \bar{t}_H$ implies $\Pi_X > 0$ and thus $V > 0$. *(ii)* follows directly from (13). For $\bar{t}_H = 0$ any $t_X \in [0, 1]$ implies $\Pi_X = 0$ and thus $V = 0$, which proves *(iii)*. **q.e.d.**

4.2 Optimization of the Large Country

For any given tax rate of the tax haven $\bar{t}_X$, the government of the large country sets its tax rate t_H to maximize welfare of its citizens $U(t_H, \bar{t}_X)$.

For certain parameter values the best response function of the large country is discontinuous. In these cases we find that the there is a threshold level of t_X depending on parameters. Below this threshold level the large country chooses a tax rate implying a strictly positive tax difference, above this level it chooses a tax difference of zero, i.e. $t_H = \bar{t}_X$. As the best response of the tax haven always implies a positive tax differential, any equilibrium has to lie below the threshold. For notational convenience we define T_3 which collects constants

$$T_3 \equiv f_t^{\epsilon}\ T_1^{-\epsilon} a_m^{\gamma-(\sigma-1)}.$$

The following proposition describes the discontinuous best response function of the large country.

Proposition 4 *Under financial integration, the implicit best response of the large country for a given* $\bar{t}_X$ *is*

$$t_H^{\rho>0} = min\left\{\frac{(\beta-1)\,(T_3 - \rho^{\epsilon})}{\epsilon\ \beta\ \rho^{\epsilon-1}}; 1\right\} \tag{14}$$

as long as the implied value of t_H is large enough to satisfy

$$\rho^{\epsilon} \geq \frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1}+(\beta-1)}\ T_3. \tag{15}$$

Otherwise the best response is given by

$$t_H^{\rho\leq 0} = \bar{t}_X. \tag{16}$$

Proof: Because this is an important result and in order to build intuition, we go into some detail proving this proposition. We first derive equations (16) and (14). They represent two different cases which we discuss separately. In case 1 the large country sets a tax rate below or equal the rate of the tax haven. In case 2 it sets a higher rate. Which of the two cases represents the best response of the large country is determined by condition (15). The derivation of this condition and the proof that (14) is indeed a welfare maximum as long as (15) holds are provided in the appendix.

Case 1: First consider the case in which the government sets a tax rate below or equal to the tax haven's rate. All firms then pay taxes at home. Total welfare in this case is:

$$U^{\rho\leq 0}(t_H,\bar{t}_X) = \bar{U} + (1-t_H)\ \Pi_H^A + \beta\ t_H\ \Pi_H^A. \tag{17}$$

Since the government values public expenditure more than expenditure on the homogenous good by its citizens ($\beta > 1$) it sets the highest possible tax rate that satisfies $t_H \leq \bar{t}_X$.[15] This tax rate is given by (16).

Case 2: Now consider the case where the government sets a tax rate that satisfies $t_H > \bar{t}_X$. Welfare in the large country is then given by

$$U^{\rho>0}(t_H,\bar{t}_X) = \bar{U} + (1-t_H)\ \Pi_H + (1-t_X)\Pi_X - N_X f_t + \beta\ t_H\ \Pi_H \tag{18}$$

Where $\Pi_H = \Pi_H^A - \Pi_X$ reflects the fact that some of the tax base in the home country flows to the tax haven when the tax differential is positive.

From the first order condition of the maximization problem, we derive (14). Where t_X and $t_H^{\rho>0}$

[15] Note that in the autarky equilibrium, the large country also sets the highest possible tax rate under the condition that $t_H^A \leq 1$.

enter in the tax differential ρ. Equation (14) provides a relatively simple implicit solution for the best response function of the government which only depends on the tax differential and parameters of the model. For the time being we assume that the second order condition for a welfare maximum is satisfied. We will see below that this is indeed the case in all relevant cases.

The best response function of the large country: For any possible value of t_X it is optimal for the large country to set its tax rate either according to (16) or to (14). The government chooses t_H according to (14) as long as $U(t_H^{\rho>0}, \bar{t}_X) \geq U(t_H^{\rho\leq 0}, \bar{t}_X)$ and according to (16) otherwise. In appendix A we show that (under the assumption that the second order condition in the derivation of (14) is satisfied) (14) is the best response iff condition (15) is satisfied. The best response is (16) otherwise. Furthermore we show that for all values of t_H that satisfy (15), equation (14) is indeed a welfare maximum. This proves Proposition 4. **q.e.d.**

The following Corollary to Proposition 4 states for which values of the tax haven's tax rate the best response of the large country is given by (14) and for which values it is given by (16).

Corollary 1 *Under financial integration when firms have the possibility to do profit shifting FDI,* (14) *is the best response of the large country to all values of the tax rate of the tax haven that satisfy*

$$t_X \leq \left((\beta - 1) \left(\frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1} + (\beta - 1)} \right)^{\frac{1-\epsilon}{\epsilon}} - (\beta - 1 + \epsilon\beta) \left(\frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1} + (\beta - 1)} \right)^{1/\epsilon} \right) \frac{T_3^{1/\epsilon}}{\epsilon\beta}. \quad (19)$$

For all values of t_X *that do not satisfy this condition,* (16) *is the best response of the large country.*

Proof: see appendix B.

While conditions (15) and (19) might look a bit complicated, it is important to note that they only depend on preference parameters, the fixed cost of becoming a multinational and the degree of firm heterogeneity.

5 Equilibrium

In this section the equilibrium of the model is derived and the effect of firm heterogeneity and market structure on existence and shape of the equilibrium is discussed. We focus on the case where governments set tax rates simultaneously. The case where the large country has a first mover advantage is discussed in the appendix C.

5.1 Equilibrium of the Tax Game

Any intersection of the two best response functions is a Nash equilibrium. As the tax haven always undercuts, (13) and (16) never intersect. Only (13) and (14) can intersect. Label t_X^x and t_H^x the tax rates at the intersection. The following Proposition states under which condition this intersection is a Nash equilibrium of the tax game and what the equilibrium tax rates are.

Proposition 5 *An intersection of* (13) *and* (14)*, is a Nash Equilibrium of the tax game if and only if for the tax rate of the tax haven at the intersection, the* t_H^x *implied by* (14) *is large enough to satisfy* (15)*. Equivalently, if* t_X^x *satisfies condition* (19)*. The equilibrium tax rates are then given by*

$$t_X^* = \frac{1}{\epsilon}\left(\frac{\beta-1}{\epsilon\beta+2\beta-1}\right)^{\frac{1}{\epsilon}}\frac{f_t}{T_1}\,a_m^{\sigma-1} \qquad \textit{and} \qquad t_H^* = \frac{\epsilon+1}{\epsilon}\left(\frac{\beta-1}{\epsilon\beta+2\beta-1}\right)^{\frac{1}{\epsilon}}\frac{f_t}{T_1}\,a_m^{\sigma-1}. \quad (20)$$

Proof: First note that equations (13) and (16) never intersect for any positive tax rate.[16] This implies that $\rho = 0$ can never be an equilibrium. The intersection of (13) and (14) is only an equilibrium of the tax game if at the intersection (14) is the best response function of the large country. This is the case if the tax rates at the intersection imply a tax difference that satisfies (15) or equivalently if t_X^x satisfies (19). When this condition is satisfied the equilibrium tax rates can be derived taking the difference of (14) and (13) and solving for the equilibrium tax difference

$$\rho^* = \left(\frac{\beta-1}{\epsilon\beta+2\beta-1}\right)^{\frac{1}{\epsilon}}\frac{f_t}{T_1}\,a_m^{\sigma-1}. \quad (21)$$

Combined with (13) the equilibrium tax rates in (20) follow directly. **q.e.d.**

We can simplify the condition for equilibrium existence as stated in the following Corollary:

Corollary 2 *An equilibrium of the tax game exists iff*

$$\beta\left(\epsilon^3 + 2\epsilon^2 - 1\right) - \epsilon^2 \geq 0. \quad (22)$$

[16] The two functions do share a common value for $t_H = t_X = 0$ but $t_X = 0$ always violates (19) so at this common point (16) does not coincide with the best response of the large country.

Proof: This follows directly from plugging (21) into condition (15) and simplifying. **q.e.d.** [17]

Note for later reference that the equilibrium tax rates do not depend on the maximum cost level. This can be seen using the definition of T_1 in equation (4). Based on these results, we can derive the equilibrium cutoff cost in equilibrium:

$$a^{**} = \left(\frac{\beta - 1}{\epsilon\beta + 2\beta - 1} \right)^{\frac{1}{\epsilon(\sigma-1)}} a_m. \tag{23}$$

The equilibrium number of firms choosing 'profit shifting' FDI is

$$N_X^* = \left(\frac{\beta - 1}{\epsilon\beta + 2\beta - 1} \right)^{\frac{\epsilon+1}{\epsilon}}. \tag{24}$$

The tax base that flows to the tax haven in equilibrium is

$$\Pi_X^* = \frac{\beta - 1}{\epsilon\beta + 2\beta - 1} \frac{\alpha}{\sigma}. \tag{25}$$

To give some intuition, the following subsection provides some graphical examples of best response functions calculated for specific parameter values. They illustrate the equilibrium for different parameter constellations (in particular $\epsilon > 1$ and $\epsilon < 1$) as well as the case when no equilibrium exists.

5.2 Cost Distribution, Market Power and Equilibrium Existence

here: Figure 1

Figure 1 illustrates the case of $\epsilon > 1$. For this parameter value condition (15) always holds. The best response function of the large country is continuous and is given by (14) for all values of t_X. It is represented by the solid line. The dashed line plots the best response of the tax haven, equation (13). As stated in Proposition 3 for a given positive $\bar{t}_H > 0$ the tax haven always undercuts the large country.

here: Figure 2 and Figure 3

[17] The result in Corollary 2 allows to narrow down the possible range for ϵ in our model. Given $\beta > 1$ we find a lower bound for equilibrium existence at about 0.618. This can be obtained by solving the equation $-1 + 2\epsilon^2 + \epsilon^3 = 0$, which represents the limit case for $\beta \to \infty$.

Figures 2 and 3 illustrate the case of $\epsilon < 1$. Now the best response of the large country is discontinuous and the equilibrium does not always exist. For low values of $\bar{t}_X$ the best response of the large country is given by (14). For high values, the optimal response of the large country is to set the same tax rate as the tax haven.

In Figure 2 the equilibrium exists. The discontinuity of the best response function of the large country is far enough to the right, so that the two best response functions intersect and an equilibrium exists. Figure 3 uses the same parameter values except for the fact that γ is lower, which implies a lower ϵ. In this case the best response function of the tax haven never intersects with the response function of the large country and thus no equilibrium exists. Condition (15) is violated: the discontinuity lies too far to the left.

To build intuition consider the best response of the large country to $\bar{t}_X = 1$. By setting $t_H = 1$ the large country maximizes tax revenue as no firms will leave the country. For a lower $\bar{t}_X$ the large country faces a trade off between extensive and intensive margin. When it sets $t_H = \bar{t}_X < 1$ it eliminates the extensive margin effect but it suffers a loss at the intensive margin as it has to lower its tax rate. We will see below that for a low value of ϵ the tax base reacts strongly to tax differences i.e. the extensive margin effect is strong. So for values of $\bar{t}_X$ sufficiently close to unity it can be optimal for the large country to prevent any outflows by setting $t_H = \bar{t}_X$. As we lower $\bar{t}_X$, however, the losses of the large country on the intensive margin become larger until the point is reached, where the large country accepts some outflows in order to tax the remaining firms at a higher rate.

The lower the shape parameter of the cost distribution and the higher the substitutability across goods (together implying a low ϵ) the earlier (i.e. for lower values of $\bar{t}_X$) the discontinuity arises. Additionally, equation (13) implies that when ϵ is low the tax haven undercuts the large country by less. Graphically, this translates into a reaction function that is closer to the diagonal. Through both effects a low ϵ makes the existence of an equilibrium less likely.

Clearly, the value of ϵ is crucial in the model. Using French firm level data, Eaton, Kortum, and Kramarz (2008) provide an estimate of ϵ of 1.46.[18] This value is larger than one so in our model the equilibrium would always exist. They do, however, also provide some evidence that points at a value of ϵ smaller unity. It is thus important not to exclude the possibility of $\epsilon < 1$ from the theoretical analysis.

[18] Using the method of simulated moments, they provide an estimate of (in our notation) $\epsilon + 1$ of 2.46 with a standard error of 0.1.

6 The Role of Industry Structure and Market Power

As discussed before all firms with a unit cost below the cutoff cost level in equation (8) opt for 'profit shifting' FDI when a positive tax difference emerges. For the tax game the distribution of profits across different cost levels is thus crucial. In this context it is especially important whether the most productive firms (the most likely to shift profits) account for a large fraction of aggregate profits. In this section we show how the distribution of firm profits is determined by the cost distribution and the extent of market power in the economy. Furthermore we show that the elasticity of the tax base also depends on these two determinants. Finally we analyze the impact of firm heterogeneity on the equilibrium outflow of tax base and on equilibrium tax rates.

6.1 Cost Distribution, Market Power and the Tax Base

In this subsection we investigate the role of the distribution of cost levels and market power on the distribution of aggregate profits (the tax base) across firms. Define $\Pi(a)$ as the share of aggregate profits accounted for by firms with cost level a: $\Pi(a) \equiv [\pi(a)g(a)]/\Pi_H^A$. It can be shown that $(\partial/\partial a)\Pi(a) = (\gamma - \sigma)(\gamma - (\sigma - 1))\ a^{\gamma-\sigma-1}$. The sign of this partial derivative thus depends on the relative size of γ and σ.

here: Figure 4

Figure 4 provides a graphical illustration of the distribution of the tax base across cost levels.[19] The thick dashed line plots the density of firms with cost level a and the solid line represents the share of profits of these firms in aggregate profits.[20] The thin vertical dashed line marks the equilibrium cutoff cost level given by (23). The area under the solid curve that is to the left of the a^{**} line thus represents the fraction of aggregate profits shifted to the tax haven. While the area under the dashed line represents the measure of firms that account for these profits: the multinationals.

The four graphs in Figure 4 illustrate how an increase in ϵ affects the distribution of the tax base. In the first three graphs we set $\sigma = 6$ and increase the Pareto parameter γ such that we obtain values of $\epsilon = 0.75$, $\epsilon = 1$ and $\epsilon = 1.5$. The latter value corresponds to the estimate of Eaton, Kortum, and Kramarz (2008). The lowest value corresponds to the value of ϵ implied in their analysis

[19] In all graphs we have set the maximum cost level a_m to unity.

[20] While an increase in the Pareto parameter γ does not affect aggregate profits, changes in the elasticity of substitution σ do. Normalizing with aggregate profits, the area under the solid line remains equal to unity in all graphs.

based on the sales of French firms in France conditioned on entry into a particular foreign market.[21] In the last graph we keep the value of $\gamma = 13.75$ and lower σ to 4 which increases ϵ to 3.58.

The four graphs illustrate how the shape parameter of the cost distribution γ and the elasticity of substitution σ affect the distribution of the tax base and the equilibrium cutoff. In the first graph firms are very heterogeneous in the sense that there is a relatively large number of firms with low cost levels. The solid line shows that these firms account for a relatively large fraction of aggregate profits. As γ rises in the next two graphs, the measure of firms with high cost levels increases. Since these firms are more numerous they also account for a larger fraction of aggregate profits. The last graph shows that a decrease in σ has a similar effect. Keeping γ constant and reducing σ to 4, the high cost firms account for an even higher share of aggregate profits. The intuition is that since consumers are less able to substitute the goods of the high cost producers, the share of profits of this group rises.

The graphs in Figure 4 illustrate the crucial role the shape of the distribution of aggregate profits plays for the mechanisms of the model. The more the distribution of firms is skewed to the left, i.e. the more of the profits is generated by low productivity firms, the less the tax base reacts to tax differences. We have seen above that the low productivity firms are the least likely to opt for profit shifting FDI. When these firms account for a large fraction of aggregate profits (which is the case when ϵ is high) a given tax difference leads to lower outflows as only the (few) productive firms are shifting profits abroad.

A low degree of firm heterogeneity thus 'protects' the large country from tax competition as it makes its tax base less reactive to differences in tax rates.[22] In order to attract tax base the tax haven has to undercut strongly to increase the cutoff cost level and to attract the firms and their profits. This mechanism is explained in more detail in the following subsection.

[21] Eaton, Kortum, and Kramarz (2008) estimate ϵ using the method of simulated moments. As mentioned above their estimate is 1.46. They provide some additional information of plausible values of ϵ exploiting the relation between ϵ and some observations in the data. In their Figure 3a they plot average sales of firms in France (conditional on entry into market n) against the number of French firms selling in market n. They find that firms that serve markets which are served by a low number of firms tend to have higher sales in France. From the slope of this relationship they deduce a value of ϵ of 0.75. A different way to obtain a value for ϵ is to use the plot of the export intensity on the number of firms selling an a particular market. The slope implies an ϵ of 1.63 which is much closer to their estimate of 1.46.

[22] We will investigate the role of firm heterogeneity using a mean preserving spread of the productivity distribution below.

6.2 The Elasticity of the Tax Base

To gain a better understanding of the role industry- and market structure play in the tax game it is of interest to look at the elasticity of the tax base with respect to the tax rate. For a given tax rate of the large country the tax haven faces an elasticity of

$$\eta^x = \epsilon \, \frac{t_X}{\bar{t}_H - t_X} = \left(\frac{\gamma}{\sigma - 1} - 1 \right) \, \frac{t_X}{\bar{t}_H - t_X}. \tag{26}$$

It is determined by three factors: the degree of firm heterogeneity, the substitutability between goods and the tax rate of the tax haven.

The elasticity increases in t_X and becomes very large when t_X is close to $\bar{t}_H$. In its revenue maximization the tax haven faces a trade off between the tax revenues per firm taxed (intensive margin) and the number of firms taxed (extensive margin). When t_X is close to $\bar{t}_H$, the elasticity is high and the extensive margin effect on tax revenues dominates. Revenue maximization implies that the tax haven lowers its tax rate until the intensive- and extensive margin effect exactly offset. This is the case when the elasticity equals unity.

How strongly the tax haven undercuts the large country thus depends on ϵ i.e. on industry structure and market power. The higher ϵ, the lower the level of t_X necessary to equate the elasticity to unity. This explains the pattern of the cutoff cost level observed in Figure 4. The case of a high γ and low σ corresponds to the last graph in Figure 4. The productive firms only account for a small part of the tax base so that the tax haven has to decrease its tax rate a lot until it reaches a point where the intensive margin effect is strong enough to offset the extensive margin effect. In equilibrium this policy results in a higher cutoff cost level. The opposite is the case in the first graph where firm heterogeneity is stronger and market power lower. In this case already a moderate degree of undercutting allows the tax haven to attract a relatively large fraction of the tax base as high productivity firms account for a large fraction of profits. As a result the tax haven undercuts less, which implies a low cutoff cost level.

It is important to note that the reason for a strong undercutting of the tax haven when ϵ is high is that it is *difficult* for it to attract tax base. A policy of strong undercutting might look aggressive when comparing equilibrium tax rates. According to the model, however, strong undercutting reveals that the tax haven is not a great danger for the tax base of the large country: the tax haven sets the low rate because its industry and market structure protect the large country from tax competition.

6.3 Heterogeneity, Equilibrium Profit Shifting and Tax Rates

In this section we investigate the role of firm heterogeneity a bit further. We have seen above that the shape parameter of the cost distribution plays a crucial role. Together with the maximum cost level a_m it determines the number of firms with different cost levels. Keeping a_m fixed at unity we first illustrate how an increase in γ affects the tax rates set, the number of multinationals and the fraction of the tax base leaving the country. We then illustrate how γ affects these variables via its impact on the degree of firm heterogeneity (as opposed to its impact on average productivity).

here: Figure 5

The first graph in Figure 5 plots the fraction of the tax base flowing to the tax haven (solid line) as a function of the shape parameter γ. The fraction of firms that choose 'profit shifting' FDI is given by the dashed line. An increase in γ implies that the tax base is less reactive to tax differences and thus the fraction of the tax base attracted by the tax haven decreases. The vertical line represents $\epsilon = 1$. For too low values of gamma no equilibrium exists. The second graph plots the equilibrium tax rates of the large country (solid line) and the rate set by the tax haven (dashed line). As γ increases, the tax base becomes less reactive to tax differences. The equilibrium tax difference increases in gamma. The reason is that the high γ 'protects' the large country from tax competition as most of the tax base is generated by small and medium sized firms. It can thus afford to set a higher tax rate. Accordingly, it is more difficult for the tax haven to attract tax base. It is forced to undercut by a larger fraction as γ increases (although t_X can be increasing for low values of γ, the tax difference is always increasing).

here: Figure 6

The correct measure of firm heterogeneity is the variance of the cost distribution. When the shape parameter γ changes, both the variance and the expectation of the cost distribution change. It is thus conceivable that the results in Figure 5 are driven by the change in the mean of the unit costs. In order to directly asses the effect of the variance, in Figure 6 we look at changes in firm heterogeneity through a mean preserving spread of the cost distribution. For each value of γ we compute the value of a_m that keeps the mean of the distribution constant. We then use these values of a_m and γ to calculate the variance and then plot it against the equilibrium tax rates and outflows in Figure 6. As expected from the reasoning above we observe the same effects.

These results imply that countries with more homogeneous industry structures are ceteris paribus less affected by international tax competition.[23]

7 Conclusions

In this paper we have proposed a stylized model of international tax competition between a large country and a tax haven. The methodological contribution is to provide a fully solvable model of international tax competition with heterogeneous firms and monopolistic competition. Our analysis reveals that firm heterogeneity and monopolistic market power are key for the tradeoffs the governments are facing. While all firms have the possibility to do 'profit shifting' FDI, only the most productive firms choose to do so.

When the very productive firms account for a large fraction of aggregate profits, the tax base reacts strongly to tax differences. This is the case when firms are very heterogeneous (relatively many relatively productive firms) and monopolistic market power is low (high degree of substitutability between goods). With such an industry- and market structure, the large country is forced to set low tax rates to limit outflows. The tax haven undercuts the large country only moderately and still attracts a large fraction of the tax base.

In the opposite case where small and medium sized firms account for a larger fraction of aggregate profits (low heterogeneity and high market power) the large country is 'protected' from international tax competition. It can set a relatively high tax rate without facing large outflows.

[23]It is clear from the analysis in the previous subsection that a country specialized in industries with low substitutability or in which the differentiated goods markets are less competitive should also be less affected by international tax competition.

References

Baldwin, Richard, and Paul Krugman, 2004, "Agglomeration, Integration and Tax Harmonisation", *European Economic Review* 48, 1–23.

Baldwin, Richard, and Toshihiro Okubo, 2006, "Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting", *Journal of Economic Geography* 6, 323–346.

Baldwin, Richard, and Toshihiro Okubo, 2008, "Tax Reform, Delocation and Heterogeneous Firms: Base Widening and Rate Lowering Reforms", CEPR Discussion Papers 6843.

Borck, Rainald, and Michael Pflueger, 2006, "Agglomeration and Tax Competition", *European Economic Review* 50, 647–668.

Bucovetsky, S., 1991, "Asymmetric Tax Competition", *Journal of Urban Economics* 30, 167–181.

Burbidge, John, Katherine Cuff, and John Leach, 2006, "Tax Competition with Heterogeneous Firms", *Journal of Public Economics* 90, 533–549.

Burbidge, J., J. DePater, G. Myers, and A. Sengupta, 1997, "A Coalition Formation Approach to Equilibrium Federations and Trading Blocs", *American Economic Review* 87, 940– 956.

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

Davies, Ronald B., and Carsten Eckel, 2007, "Tax Competition for Heterogeneous Firms with Endogenous Entry", University of Oregon Working Paper 2007-6.

Desai, Mihir A., C. Fritz Foley, and James R. Hines, 2006, "The Demand for Tax Haven Operations", *Journal of Public Economics* 90, 513–531.

Eaton, J., S. Kortum, and F. Kramarz, 2008, "An Anatomy of International Trade: Evidence from French Firms", NBER Working Paper 14610.

Fuest, C., B. Huber, and J. Mintz, 2005, "Capital Mobility and Tax Competition: a Survey", *Foundations and Trends in Microeconomics* 1, 1–62.

Hamada, K., 1966, "Strategic Aspects of Taxation on Foreign Investment Income", *Quarterly Journal of Economics* 80, 361–375.

Krugman, Paul, 1980, "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review* 70, 950–59.

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

Ottaviano, Gianmarco I.P., and Tanguy van Ypersele, 2005, "Market Size and Tax Competition", *Journal of International Economics* 67, 25–46.

Sato, Yasuhiro, and Jacques-François Thisse, 2007, "Competing for Capital when Labor is Heterogeneous", *European Economic Review* 51, 2054–2079.

Wilson, John Douglas, 1986, "A Theory of Interregional Tax Competition", *Journal of Urban Economics* 19, 296–315.

Wilson, John Douglas, 1999, "Theories of Tax Competition", *National Tax Journal* 52, 269–304.

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.

Appendix

A Proof of Proposition 4

To complete the proof of Proposition 4 in the main text, we proceed in two steps. We first derive condition (15) under the *assumption* that the second order condition for a welfare maximum holds for (14). In a second step we show that this is the case whenever (14) is the countries' best response function derived in step 1.

Step 1: To see *(i)*, note that the large country will set t_H such that $\rho > 0$ (i.e. according to (14)) as long as $U(t_H^{\rho>0}, \bar{t}_X) \geq U(t_H^{\rho\leq 0}, \bar{t}_X)$. Plugging in (18) and (17) into this condition, rearranging and using (16) and $(1 - t_X) - (1 - t_H^{\rho>0}) = \rho$ we get:

$$\rho\ (\beta - 1)\ \Pi_H^A \geq \left(\beta\ t_H^{\rho>0} - \rho\right)\ \Pi_X + N_X\ f_t. \tag{27}$$

Now first using $f_t N_x = \Pi_X\ \rho\ \frac{\epsilon}{\epsilon+1}$, then $\Pi_H^A = T_2\ T_1^{-\epsilon}\ a_m^{1-\sigma}$ and $\Pi_X = \rho^\epsilon\ T_2\ f_t^{-\epsilon}\ a_m^{-\gamma}$, simplifying and solving for $t_H^{\rho>0}$ gives:

$$t_H^{\rho>0} \leq \frac{(\beta - 1)T_3 + \rho^\varepsilon \frac{1}{\varepsilon+1}}{\beta\ \rho^{\varepsilon-1}}. \tag{28}$$

To see for which values of ρ the tax rate in (14) satisfies condition (28), we plug (14) into (28), which gives:

$$\rho^\epsilon \geq \frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1} + (\beta - 1)}\ T_3.$$

As long as for a given $\bar{t}_X$ the t_H implied by (14) is high enough to satisfy this condition, we will have $U(t_H^{\rho>0}, \bar{t}_X) \geq U(t_H^{\rho\leq 0}, \bar{t}_X)$ and thus the best response function of the large country given by (14). When the above condition is violated, we have $U(t_H^{\rho>0}, \bar{t}_X) < U(t_H^{\rho\leq 0}, \bar{t}_X)$. In this case the best response of the large country is given by (16) because it maximizes $U(t_H^{\rho\leq 0}, \bar{t}_X)$.

Step 2: It remains to be shown that for any given $\bar{t}_X$ equation (14) is a welfare maximum (and not a minimum) for all relevant values of t_H. Relevant values are all values of t_H that satisfy (15) for a given $\bar{t}_X$.[24]

First note that the second derivative of the welfare function with respect to the tax rate of the

[24] All values that do not satisfy (15) are irrelevant as in these cases the best response of the large country is given by the 'case 1' best response (16) anyway.

large country is given by:

$$\frac{\partial^2 U}{\partial t_H^2} = -2(\beta-1)\frac{\partial \Pi_X}{\partial t_H} - \frac{\partial^2 \Pi_X}{\partial^2 t_H}[(\beta-1)t_H + t_X] - \frac{\partial^2 N_x}{\partial^2 t_H} f_t.$$

For notational convenience, we define $T_5 \equiv T_2 \epsilon \rho^{\epsilon-1} f_t^{-\epsilon} a_m^{-\gamma} > 0$. We then use $\frac{\partial \Pi_X}{\partial t_H} = T_5$ together with $\frac{\partial^2 \Pi_X}{\partial t_H^2} = \frac{\epsilon-1}{\rho} T_5$ and $\frac{\partial^2 N_x}{\partial t_H^2} f_t = \epsilon T_5$ to get:

$$\begin{aligned}
\frac{\partial^2 U}{\partial t_H^2} &= -T_5 \left(2(\beta-1) + \epsilon + \frac{\epsilon-1}{\rho}\,[(\beta-1)t_H + t_X] \right) \\
&= -T_5 \left(2(\beta-1) + \epsilon - (\epsilon-1) + (\epsilon-1)\ \beta\ \frac{t_H}{\rho} \right)
\end{aligned}$$

From (14) it follows that $\frac{t_H}{\rho} = \frac{(\beta-1)\ T_3}{\epsilon\ \beta\ \rho^\epsilon} - \frac{(\beta-1)}{\epsilon\ \beta}$ so that:

$$\frac{\partial^2 U}{\partial t_H^2} = -T_5 \left(2(\beta-1) + \epsilon + \frac{(\beta-1)(\epsilon-1)}{\epsilon} \left(\frac{T_3}{\rho^\epsilon} - 1 \right) - \epsilon + 1 \right). \tag{29}$$

We need to show that equation (15) is a sufficient condition for (29) to be negative. To do so, we proceed in two steps. We first show that it is negative when the above condition holds with equality. We then show that this is also true for larger values of ρ.

Define ρ^* as the value of ρ, where (15) holds with equality. We will first determine the sign of (29) for ρ^* i.e. $\frac{\partial^2 U}{\partial t_H^2} \mid_{\rho^*}$.

$$\frac{\partial^2 U}{\partial t_H^2} \mid_{\rho^*} = -T_5 \left(2(\beta-1) + 1 - T_j - \frac{(\beta-1)(\epsilon-1)}{\epsilon} \right) \tag{30}$$

With $T_j \equiv \frac{(\beta-1)(1-\epsilon)}{\epsilon} \frac{\frac{\epsilon}{\epsilon+1} + (\beta-1)}{(\beta-1)(1-\epsilon)} = \frac{1}{\epsilon+1} + \frac{\beta-1}{\epsilon} > 0$

Simplifying and recalling that $T_5 > 0$ then gives:

$$\frac{\partial^2 U}{\partial t_H^2} \mid_{\rho^*} \quad = \quad -T_5 \left((\beta-1) + \frac{\epsilon}{\epsilon+1} \right) < 0. \tag{31}$$

This implies that for $\rho = \rho^*$ the second order condition holds and (14) is indeed the optimal response.

To see that this is true for all $\rho \geq \rho^*$, note that any value of $\rho \geq \rho^*$ can be written as $\rho = x\ \rho^*$ with $x \geq 1$. In order to obtain $\frac{\partial^2 U}{\partial t_H^2} \mid_{\rho^*}$, we have plugged in ρ^* into (29). Now considering any value of $\rho \geq \rho^*$, we can plug $\rho = x\ \rho^*$ into to (29).

It can be seen in (29) that ρ enters twice in the expression for $\frac{\partial^2 U}{\partial t_H^2}$. Entering via T_5 it does not affect the sign. To see the effect of a higher ρ on the second term, note that when we use $\rho = x\ \rho^*$, T_j in equation (30), is replaced by $T_j\ \frac{1}{x^{\varepsilon}} < T_j$. The positive effect of T_j on the sign of $\frac{\partial^2 U}{\partial t_H^2}$ is thus dampened for a any $\rho > \rho^*$. This shows that (15) is indeed a sufficient condition for (14) to be a utility maximum. **q.e.d.**

B Proof of Corollary 1

Define ρ^{jump} as the tax differential just before the regime switch to $\rho = 0$. For ρ^{jump}, (15) holds with equality. Note that $t_X^{jump} = t_H^{\rho>0}[\rho^{jump}] - \rho^{jump}$ combining this with (14) and (15), the value of t_X for which the best response for the large country switches from $\rho > 0$ to $\rho = 0$ is given by

$$t_X^{jump} = \frac{(\beta-1)\ T_3}{\epsilon\beta\ \left(\frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1}+(\beta-1)}\ T_3\right)^{\frac{\epsilon-1}{\epsilon}}} - \frac{(\beta-1)+\epsilon\beta}{\epsilon\beta}\left(\frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1}+(\beta-1)}\ T_3\right)^{1/\epsilon}.$$

This can be simplified to

$$t_X^{jump} = \left((\beta-1)\left(\frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1}+(\beta-1)}\right)^{\frac{1-\epsilon}{\epsilon}} - (\beta-1+\epsilon\beta)\left(\frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1}+(\beta-1)}\right)^{1/\epsilon}\right)\frac{T_3^{1/\epsilon}}{\epsilon\beta}.$$

To prove the inequality in Corollary 1, it remains to be shown that all values of t_X below t_X^{jump} imply that the large country sets its tax rate such that $\rho > 0$.

We know that the utility of the best response with $\rho > 0$ dominates for $t_X = t_X^{jump}$. A sufficient condition for this to hold for $t_X \leq t_X^{jump}$ as well, is that ρ decreases in $t_X\ \forall\ t_X \leq t_X^{jump}$. From before we have:

$$t_H = \frac{(\beta-1)(T_3-\rho^{\epsilon})}{\epsilon\beta\rho^{\epsilon-1}}$$

subtracting t_X on both sides and multiplying by $\epsilon\beta\rho^{\epsilon-1}$ we get:

$$\epsilon\beta\rho^{\epsilon} = (\beta-1)T_3 - (\beta-1)\rho^{\epsilon} - t_X\epsilon\beta\rho^{\epsilon-1}$$

This can be rewritten as:

$$t_X = Q(\rho) = \frac{1}{\epsilon\beta}[(\beta-1)T_3\rho^{1-\epsilon} - (\epsilon\beta+\beta-1)\rho]$$

It remains to show that $Q'(\rho) < 0\ \forall\ t_X \leq t_X^{jump}$

$$Q'(\rho) = (1-\epsilon)(\beta-1)T_3\rho^{-\epsilon} - (\epsilon\beta+\beta-1) < 0$$

This condition can be rewritten to:

$$\rho^{\epsilon} > \frac{(1-\epsilon)(\beta-1)}{(\epsilon\beta+\beta-1)}\,T_3$$

Condition (15) gives a lower bound for ρ. Plugging in this bound into the previous condition delivers:

$$\frac{(1-\epsilon)(\beta-1)}{\frac{\epsilon}{\epsilon+1}+\beta-1}\,T_3 > \frac{(1-\epsilon)(\beta-1)}{(\epsilon\beta+\beta-1)}\,T_3$$

which can be simplified to:

$$\beta > \frac{1}{\epsilon+1}$$

which is always true. **q.e.d.**

C The Stackelberg Case

In the main text we have analyzed the Nash equilibrium of the tax game. We have focused on the Nash equilibrium as it is not clear to us why a country should be a Stackelberg leader in this game. To complement our analysis of the Nash equilibrium in the main text, we now consider the tax game with Stackelberg leadership of the large country.[25] This can be done relatively easily due to the simple best response function obtained for the tax haven.

The aim is to compare the Nash and the Stackelberg equilibrium. For this comparison we can normalize $a_m \equiv 1$. It turns out that the two tax rates are strategic complements and that both countries set higher tax rates and are both better off than in the Nash equilibrium.

The Stackelberg Equilibrium: Recall that welfare of the large country is given by:

$$U = \bar{U} + \Pi_H + (\beta-1)t_H\Pi_H + (1-t_X)\Pi_X - N_X f_t.$$

When setting its tax rate, the large country takes into account the best response of the tax haven given by equation (13). The best response implies $\rho = t_H E$ where we have defined $E \equiv \frac{\epsilon}{\epsilon+1}$. Using the best response of the tax haven we can write:

$$\Pi_H = T_2(T_1^{-\epsilon} - t_h^{\epsilon}E^{\epsilon}f_t^{-\epsilon}) \qquad \Pi_X = t_H^{\epsilon}E^{\epsilon}T_2 f_t^{-\epsilon} \qquad N_X = t_H^{\epsilon+1}E^{\epsilon+2}T_2 f_t^{-\epsilon-1}$$

[25]To avoid significant analytical complexity, we refrain from analyzing the reverse case i.e. the case with the tax haven being the Stackelberg leader.

The partial derivatives are given by:

$$\frac{\partial \Pi_H}{\partial t_H} = -\epsilon t_H^{\epsilon-1} E^{\epsilon} T_2 f_t^{-\epsilon} \qquad \frac{\partial \Pi_X}{\partial t_H} = -\frac{\partial \Pi_H}{\partial t_H} \qquad \frac{\partial N_X}{\partial t_H} = \epsilon t_H^{\epsilon} E^{\epsilon+1} T_2 f_t^{-\epsilon-1}$$

and $\frac{\partial t_X}{\partial t_H} = \frac{1}{\epsilon+1}$. The partial derivative of welfare with respect to the large countries tax rate is:

$$\frac{\partial U}{\partial t_H} = (\beta-1)\Pi_H + t_H \left[(\beta-1) + \frac{1}{\epsilon+1}\right] \frac{\partial \Pi_H}{\partial t_H} - \frac{1}{\epsilon+1}\Pi_X - \frac{\partial N_X}{\partial t_H} f_t.$$

Plugging in delivers:

$$\frac{\partial U}{\partial t_H} = (\beta-1)T_2 T_1^{\epsilon} - (\beta-1)t_h^{\epsilon} E^{\epsilon} f_t^{-\epsilon} T_2 - t_h[(\beta-1) + \frac{1}{\epsilon+1}](\epsilon t_h^{\epsilon-1} E^{\epsilon} f_t^{-\epsilon} T_2)$$
$$-\frac{1}{\epsilon+1} t_h^{\epsilon} E^{\epsilon} f_t^{-\epsilon} T_2 - (\epsilon+1) t_h^{\epsilon} E^{\epsilon+1} T_1^{\epsilon+1} f_t^{-\epsilon}.$$

Which can be simplified to:

$$\frac{\partial U}{\partial t_H} = T_2 f_t^{-\epsilon} \left[(\beta-1)\left(\frac{f_t}{T_1}\right)^{\epsilon} - t_h^{\epsilon} E^{\epsilon}((\epsilon+1)(\beta-1) + 1 + \frac{\epsilon^2}{\epsilon+1})\right].$$

At this point it is convenient to check the second order condition. Differentiation delivers:

$$\frac{\partial^2 U}{\partial^2 t_H} = T_2 f_t^{-\epsilon} \left[-\epsilon t_h^{\epsilon-1} E^{\epsilon}((\epsilon+1)(\beta-1) + 1 + \frac{\epsilon^2}{\epsilon+1})\right] < 0.$$

Setting the first derivative to zero and rearranging gives:

$$t_h^{st} = \frac{\epsilon+1}{\epsilon} \left(\frac{\beta-1}{(\epsilon+1)(\beta-1) + 1 + \frac{\epsilon^2}{\epsilon+1}}\right)^{\frac{1}{\epsilon}} \frac{f_t}{T_1}.$$

This is the optimal tax rate of the large country as the Stackelberg leader. The tax haven then sets:

$$t_X^{st} = \frac{1}{\epsilon} \left(\frac{\beta-1}{(\epsilon+1)(\beta-1) + 1 + \frac{\epsilon^2}{\epsilon+1}}\right)^{\frac{1}{\epsilon}} \frac{f_t}{T_1}.$$

Stackelberg vs. Nash: In the following we compare this result to the Nash Equilibrium derived in the main text. It turns out that the two tax rates are strategic complements and that both countries are better off in the Stackelberg case. The following Proposition summarizes the results of this comparison.

Proposition 6 *In the Stackelberg equilibrium,* *(i) the large country and the tax haven both set higher tax rates than in the Nash equilibrium. I.e.* $t_H^{st} > t_H^*$ *and* $t_X^{st} > t_X^*$.

(ii) the cutoff cost level is higher, the measure of multinationals is larger and a larger share of profits is shifted abroad than in the Nash equilibrium. I.e. $a_{st}^* > a^*$, $N_{X,st}^* > N_X^*$ *and* $\Pi_{X,st}^* > \Pi_X^*$.

(iii) the large country and the tax haven both have higher levels of welfare than in the Nash equilibrium.

Proof: The equilibrium tax rates in the unique Nash equilibrium are given by equation (20). To show *(i)* we have to show that:

$$\epsilon\beta + 2\beta - 1 > (\epsilon + 1)(\beta - 1) + 1 + \frac{\epsilon^2}{\epsilon + 1}.$$

This can be simplified to: $\beta(\epsilon + 1) > 1$ which is always true. Thus $t_h^{st} > t_H^*$ and $t_X^{st} > t_X^*$.
(ii) can be established using the above result and noting that:

$$a_{st}^{**} = \left(\frac{\beta - 1}{(\epsilon + 1)(\beta - 1) + 1 + \frac{\epsilon^2}{\epsilon+1}} \right)^{\frac{1}{\epsilon(\sigma-1)}} > a^{**}$$

$$N_{X,st}^* = \left(\frac{\beta - 1}{(\epsilon + 1)(\beta - 1) + 1 + \frac{\epsilon^2}{\epsilon+1}} \right)^{\frac{\epsilon+1}{\epsilon}} > N_X^*$$

$$\Pi_{X,st}^* = \frac{\alpha}{\sigma} \left(\frac{\beta - 1}{(\epsilon + 1)(\beta - 1) + 1 + \frac{\epsilon^2}{\epsilon+1}} \right) > \Pi_X^*.$$

To prove *(iii)* note that for the large country the Nash equilibrium is a point on the best response function of the tax haven. Thus it is part of the choice set of the large country in the Stackelberg case. The First and Second order conditions deliver a unique maximum. This unique maximum is different to the Nash Equilibrium. Thus $U_{st} > U_{nash}$. For the tax haven, note that $V_{st} = t_X^{st}\Pi_X^{st}$. From *(i)* and *(ii)* it follows that $V_{st} > V^*$. **q.e.d.**

D Figures

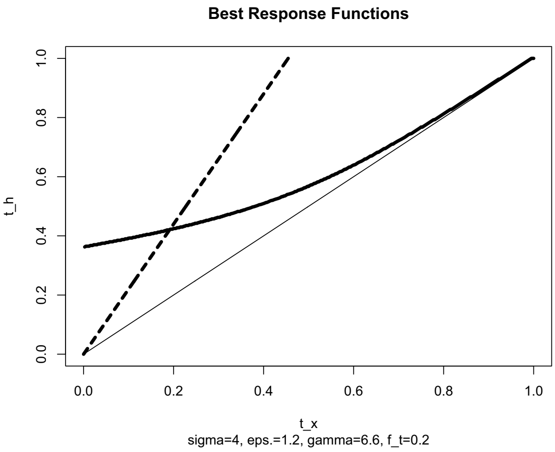

Figure 1: This Figure provides a numerical example for the equilibrium of the tax game for $\epsilon > 1$. In this case the best response function of the large country is continuous (bold solid line) and there exists a unique intersection with the best response function of the tax haven (dashed line). The best response of the large country implies $\rho > 0$ for all tax rates of the tax haven. The parameter values chosen are $\sigma = 4$, $\gamma = 6.6$ (implying $\epsilon = 1.2$), $a_m = 1$ and $f_t = 0.2$.

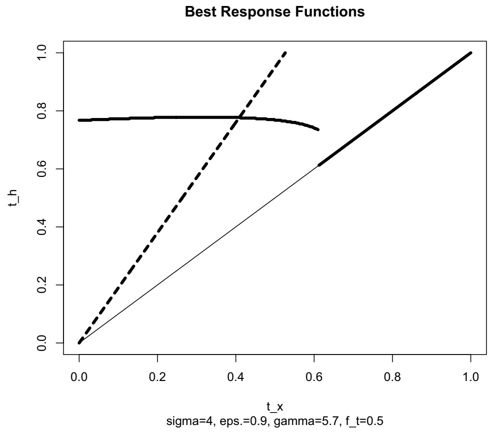

Figure 2: This numerical example illustrates that for $\epsilon < 1$ the reaction function of the large country (bold solid line) is discontinuous. For low values of t_X the large country sets a higher tax rate to finance public expenditure accepting an outflow of tax base. A low ϵ implies, however, that the most productive firms account for a large fraction of the tax base. Setting the tax differences to zero (keeping the most productive firms paying taxes at home) is optimal for high values of t_X. In this example with $\epsilon = 0.9$ the equilibrium exists. The parameter values chosen are $\sigma = 4$, $\gamma = 5.7$, $a_m = 1$ and $f_t = 0.5$.

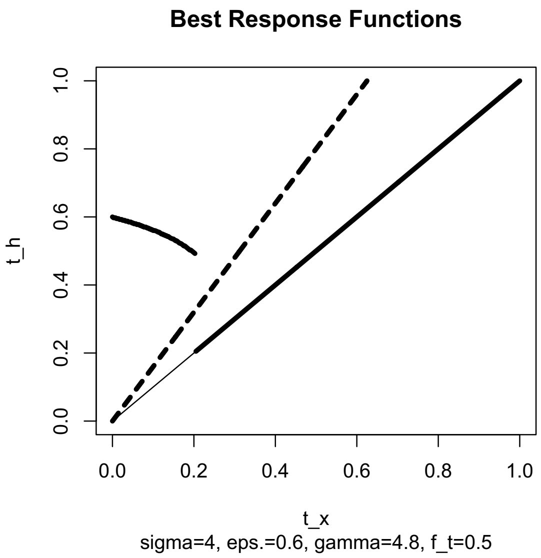

Figure 3: This numerical example uses the same parameter values as Figure 2 except that γ (and thus ϵ) is lower. This implies that the most productive firms account for a large fraction of the tax base and that the tax base thus reacts strongly to tax differences. Trying to avoid these firms from paying taxes abroad, the large country sets the tax difference to zero already for low levels of t_X. Since the tax haven always undercuts, this cannot be an equilibrium. In this example no equilibrium exists. The parameter values chosen are $\sigma = 4$, $\gamma = 4.8$ (implying $\epsilon = 0.6$), $a_m = 1$ and $f_t = 0.5$.

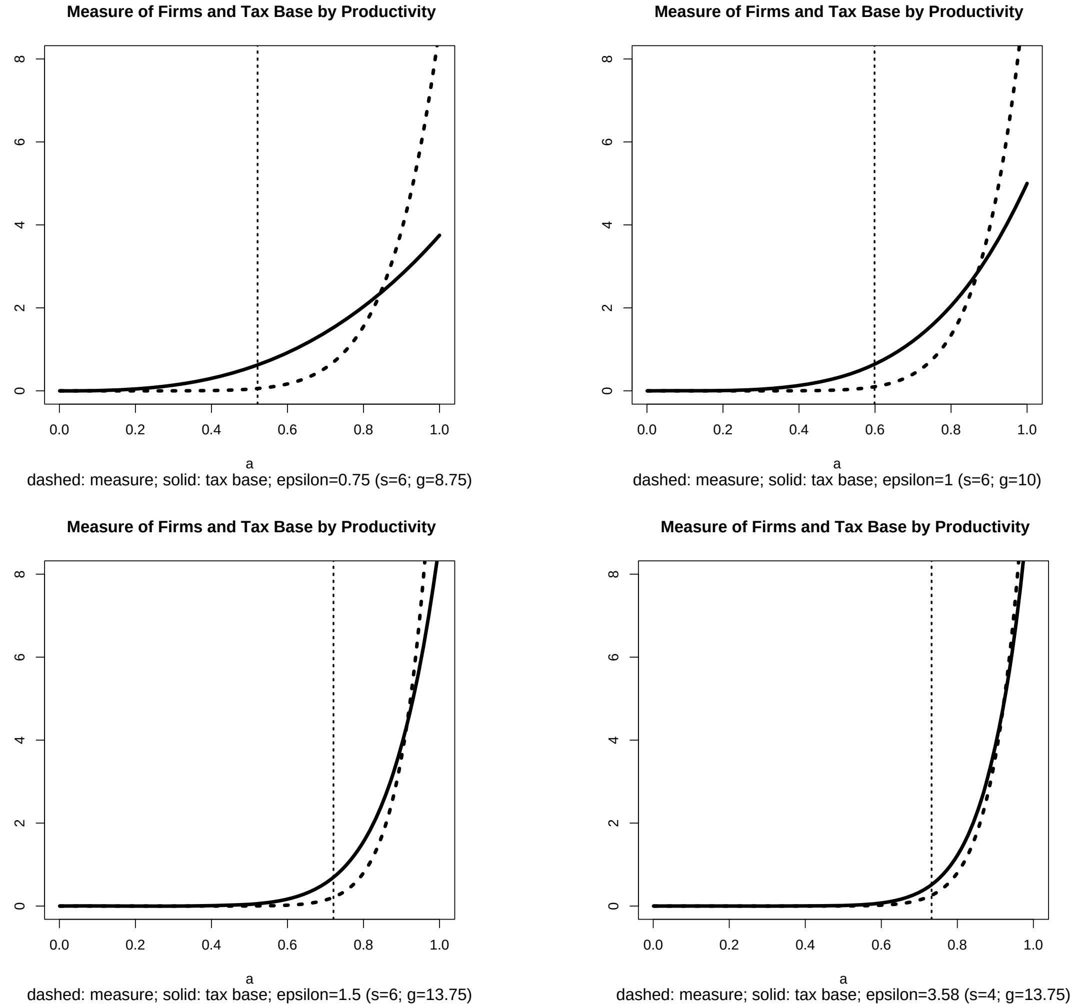

Figure 4: These three graphs illustrate the effect of the Pareto parameter γ and the elasticity of substitution σ on the distribution of the tax base (solid line). Profits are normalized by overall aggregate profits (Π_H^A) thus the area under the curve is unity in all cases. A point on the curve represents the share of overall profits firms with cost level a account for. The dashed line plots the density of firms $g(a)$. The dashed vertical line plots the equilibrium cutoff level. We use $a_m = 1$. In the first three graphs we set $\sigma = 6$ and vary γ such that ϵ rises from 0.75 ($\gamma = 8.75$) to 1 ($\gamma = 10$) and then to 1.5 ($\gamma = 13.75$). In the last graph we keep $\gamma = 13.75$ and decrease σ to 4 which implies $\epsilon = 3.58$.

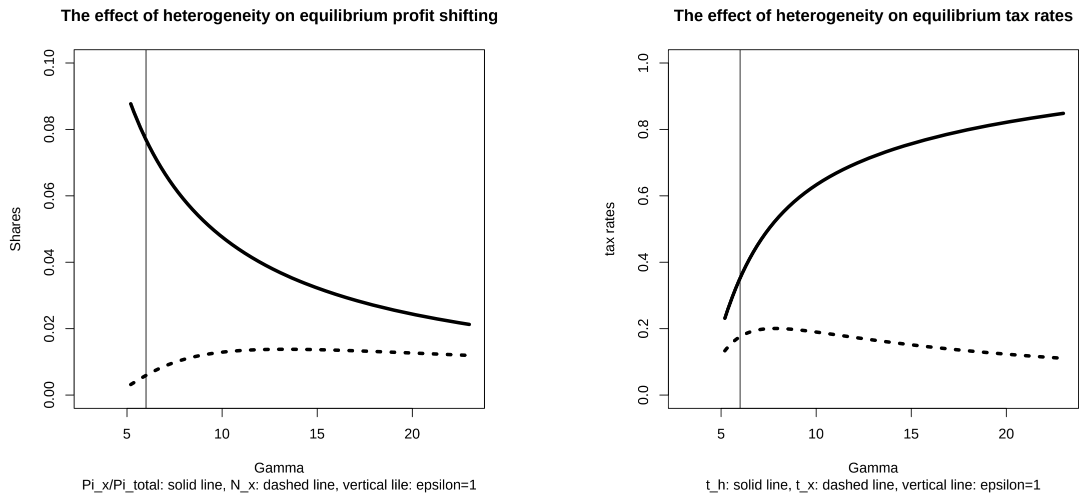

Figure 5: In the first graph the solid line represents the fraction of the tax base flowing to the tax haven as a function of γ. The fraction of firms that choose 'profit shifting' FDI is given by the dashed line. The second graph plots the equilibrium tax rates of the large country (solid line) and the rate set by the tax haven (dashed line). In both graphs the vertical line represents $\epsilon = 1$. For too low values of gamma no equilibrium exists.

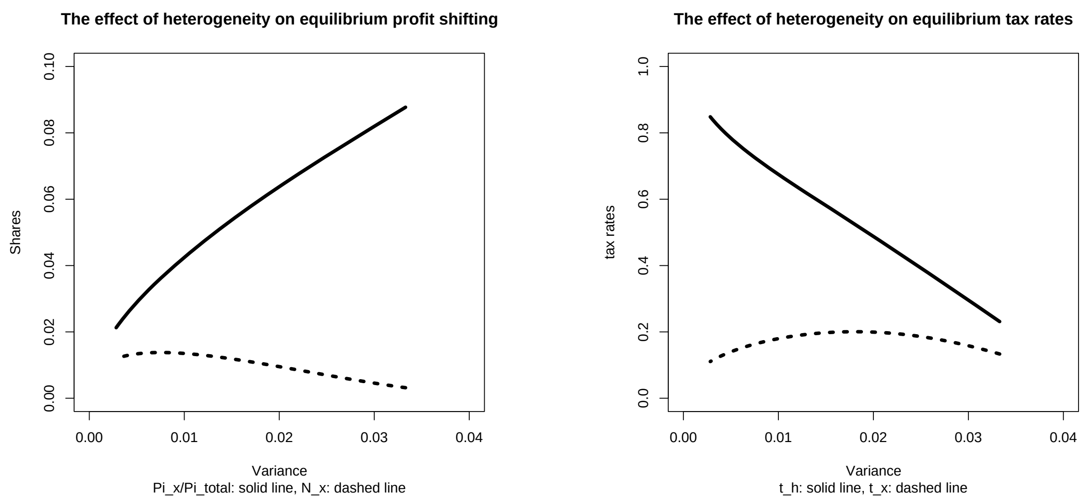

Figure 6: These graphs plot the same variables as in Figure 5 against the variance of the cost distribution which is a direct measure of firm heterogeneity. We vary γ and a_m simultaneously to generate a mean preserving spread. In line with the results in Figure 5 outflows increase and the tax difference decreases when firms are more heterogeneous.