European University Institute
DEPARTMENT OF ECONOMICS

Essays in Public Finance

Simon Skipka

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, 03 December 2019

European University Institute
Department of Economics

Essays in Public Finance

Simon Skipka

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Andrea Galeotti, London Business School, Supervisor
Prof. Dominik Sachs, University of Munich
Prof. Dominika Langenmayr, Catholic University Eichstätt-Ingolstadt
Prof. Kai Konrad, Max Planck Institute for Tax Law and Public Finance

European University Institute

DEPARTMENT OF ECONOMICS

Researcher declaration to accompany the submission of written work

I Simon Skipka certify that I am the author of the work Essays in Public Finance I have presented for examination for the PhD thesis at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the *Code of Ethics in Academic Research* issued by the European University Institute (IUE 332/2/10 (CA 297).

I confirm that chapter 2 was jointly co-authored with Mr Rafael Barbosa and I contributed 50% of the work.

Signature and Date:

Simon Skipka

November 25, 2019

Abstract

This thesis consists of three independent essays in public finance. While differing in methodology and scope, the essays are unified by their topicality in the policy discussion

The first chapter discusses the interaction between tax competition and the intensity of profit shifting, modelling heterogeneity between countries and across firms. The project is motivated by recent policies that allow for a limited degree of profit shifting in order to mitigate production inefficiencies, which are a common side-effect of profit shifting restrictions. By modelling strategic effects in a general equilibrium model of tax competition, my model challenges the effectiveness of policies with limited profit shifting. In a second contribution, I show that marginal profits from allowing for limited profit shifting are concentrated on firms with an intermediate degree of mobility. The concentration result is new to the literature and could explain the introduction of limited profit shifting policies in settings, in which the headline rationale is incapable of.

The second chapter, co-authored with Rafael Barbosa, addresses the puzzle that, despite its theoretical merits, Land Value Taxation (LVT) is not a common policy instrument in most countries. One of the main reasons is uncertainty regarding its distributional impacts. This uncertainty has not been settled by the literature, due to a lack of appropriate data. We overcome this obstacle by the construction of a unique household level dataset for a sample of German homeowners in 2017 and use the data to study the distributional effects of implementing a LVT compared to a standard property tax. We find a LVT to be equally progressive if implemented at the federal level, but less progressive if implemented at the regional level. Quantitatively a revenue-neutral reform from a standard property tax to a LVT would increase the average tax burden of the lowest income quintile of homeowners by 25%.

In the third chapter I discuss a pattern in local business tax rates, which so far has been largely overlooked by the literature: Differences in tax rates between neighboring municipalities of similar size, Twin Tax Differences (TTD). First, I show this pattern to be a significant phenomenon in local business taxation: About 35% of the 11,000 German municipalities have a TTD of at least two percentage points. Second, I provide a novel theory for the emergence of TTD as an outcome of tax competition between adjacent, fully homogeneous municipalities. Based on the analysis, I argue how TTD, and the mechanisms behind, should impact public policies, both in terms of redistribution and incentive provision.

Acknowledgements

Many people have been important to me during the time of my doctoral studies. In the acknowledgements to this thesis, I will only mention those, who had a direct impact on my work. All the people and moments, who made my last four years socially unforgettable will be separately mentioned in a post on social media in the next days.

I thank Andrea Galeotti for all the time he spent on my projects. His dedication to my work, his incomprehensible intellectual capacity and his honesty made him the most important pillar to rely on throughout the work on my dissertation. I thank Dominik Sachs for his commitment to me and my research. He always asked the right questions to let me remember the bigger picture and constantly stimulated me to think outside my set of equations. Moreover, I thank both of my supervisors for the conviction they had in me and the opportunities they gave me based on that.

My work has benefitted enormously from discussions with Nina Bobkova, Giacomo Calzolari, Denis Gorea, Piero Gottardi, Andreas Haufler and Niels Johannesen.

During the time of my dissertation, I have received immeasurable support from the administration at the economics department.

My fellow researchers at the EUI have been a constant source of feedback and support. While I am indebted to many more of them, I take this opportunity to mention only a few. I thank Rafael Barbosa for being an intellectually challenging, hard-working co-author and for being a great partner to discuss papers/seminars of all sorts. I thank Benedikt Dengler for stimulating my interest in geo-data and for showing me the way to the Ministry. I thank David Koll for being the first to test my ideas/results on and for carefully preparing each of my external talks together. I thank Essi Kujansuu for giving the most critical feedback on my written pieces and for being the most valuable part of the secret network working group. I thank Yorgos Manalis for all his feedback on my work and for the organization of numerous, repetitive talks of mine in the informal working group. I thank Chiara Santantonio for boosting the quality of my drafts with her cleverness and for getting crazy about my comma placement.

Ich danke meinen Großeltern, Christa und Christoph, Ingeborg und Wolfgang, für die liebevolle Unterstützung und das bedingungslose Vertrauen. Ich danke meinen Eltern, Eveline und Christoph, für alles. Ich bin unfassbar dankbar und glücklich, euch zu haben.

Contents

Part I

Tax Competition with Limited Profit Shifting

Tax Competition with Limited Profit Shifting

Simon Skipka*

Abstract

This paper discusses the interaction between tax competition and the intensity of profit shifting, taking into account heterogeneity between countries and across firms. The project is motivated by recent policies that allow for a limited degree of profit shifting in order to mitigate production inefficiencies, which are a common side-effect of profit shifting restrictions. By modelling strategic effects, my model challenges the effectiveness of policies with limited profit shifting. Increasing the scope for profit shifting decreases the incentive for high productive countries to compete for mobile tax revenues, because they cannot rely on their competitive advantage in productivity anymore to 'lock-in' tax revenues of multinational enterprises. As a strategic response, those countries increase their tax rates to extract higher tax revenues from their immobile tax base. To avoid the increasing level of taxation in high productive countries, firms allocate more production to low productive countries and so, raise the production inefficiency. In a second contribution, my model shows that marginal profits from allowing for limited profit shifting are concentrated on firms with an intermediate degree of international mobility. In order to gain from more flexibility in shifting profits, a firm needs to have some investment in low tax countries *and* needs to generate a production inefficiency from that investment. Immobile firms fail on the first, highly mobile firms on the second condition.

Keywords: Tax Competition, Profit Shifting, Apportionment Regimes.

JEL Classification: H25, H26, H73

*Department of Economics, European University Institute, Florence, Italy (email: simon.skipka@eui.eu). I would like to thank Andrea Galeotti, Andreas Haufler, Niels Johannesen, Dominik Sachs as well as seminar participants in Bochum, Florence, Munich (LMU and MPI) and Oxford for valuable comments and suggestions.

1 Introduction

I consider *Tax Competition* as a situation in which countries offer a business environment to firms, for example infrastructure or human capital, and charge corporate income tax rates in exchange, as the price of using the business environment. Firms are potentially free to choose where to locate from a set of countries, so that the tax setting of a country is constrained by the tax setting of other countries. The resulting interdependence of tax rates finally leads to a situation of Tax Competition for tax revenues and investment. At the same time, *Profit Shifting* is considered as the effort of Multinational Enterprises to claim profits in low tax countries, without performing a respective share of economic activity there[1]. This project will be precise about the definition of a *respective share* of economic activity, however, it will be agnostic about the means of profit shifting.

While the welfare effects of tax competition are ambiguous, profit shifting has been identified as a threat to sustainable public finances[2]. Accordingly, the containment of profit shifting was high up on the agenda of policymakers during the last decades. I take this agenda as given and contribute to the question of *how* to contain it. In particular, I consider policies that seek to establish a link between a firm's physical presence and the ability to claim profits in a country. For decades, such policies have been promoted as an effective tool to curb artificial profit shifting. However, optimal implementation and the overall impact on the economy are still actively debated in academia and policy institutions. Motivated by recent policy-advances, I approach the discussion from a slightly different perspective. I focus on policies that seek to establish a link between physical presence and the ability to claim profits and can be seen as intermediate solutions between two classic benchmarks.

The two benchmarks are: (a) unregulated profit shifting - the fraction of profits taxable in a country is independent of the fraction of production (b) Proportional Apportionment - the fraction of profits taxable in a country equals the fraction of production in that country[3]. This paper will discuss intermediate policies, which are defined in the following way: The maximum fraction of profits a firm is allowed to claim in a country depends on the fraction of production, but overproportionately. I consider such principle as a situation with *Limited Profit Shifting* and will model it in line with a recent prominent policy, the OECD's modified Nexus-approach. Using the OECD's framework provides an immediate application for this paper's theoretical results. Furthermore, the functional form of the modified Nexus-approach nests the classic benchmarks as limit cases.

The official rationale for the introduction of limited profit shifting policies is production efficiency. In an economy with unregulated profit shifting, the firms' decision on the international allocation of production is independent of tax-based incentives. Irrespective of the place of production, in the end of the financial year a firm is free to claim profits in the country with the lowest tax rates. Thus, unregulated profit shifting implies an efficient international allocation of production. By contrast, the literature on Proportional Apportionment has shown, both empirically and theoretically, that such regimes trigger an inefficiency in the allocation of production[4]. Firms have an incentive to increase production in low tax countries above the production efficient

[1]For a more general discussion on the definition of tax competition, see Devereux and Loretz (2013).

[2]Estimates on the revenue losses implied by profit shifting greatly differ due to methodological ambiguities. For a recent survey of the literature, see Riedel (2018).

[3]In reality, profits are apportioned with respect to a wide class of factors like production, costs, sales. For the rest of this project, I focus on production-based apportionment only.

[4]For an example of a recent empirical paper, see Eichfelder et al. (2018), the seminal theoretical paper on this matter is Gordon and Wilson (1986).

level in order to reduce their global tax bill. In consequence, the containment of profit shifting seems to come at the expense of a decrease in production efficiency. More concretely, regulating profit shifting is perceived to face a *containment-efficiency* trade-off.

However, to date the existence of such trade-off is mostly derived from a comparison of the two benchmarks. The first contribution of my paper expands the analysis by asking the following Research Question: 'Do intermediate policies unambiguously increase production efficiency compared to Proportional Apportionment?'. If the answer were 'Yes', the introduction of policies like the modified Nexus (instead of a standard Proportional Apportionment) could be interpreted as a shift in the preferences of a supranational policymaker, leaning more towards global production efficiency. However, the answer is that it may, in my model will, be the case that starting from Proportional Apportionment and allowing for some limited degree of profit shifting initially increases the inefficiency even more. Only if the limitations on profit shifting are already sufficiently weak, a further lift decreases the inefficiency, until the production efficient benchmark of unregulated profit shifting is reached.

To understand the intuition of the result, consider a sketch of the model: The model hosts two types of countries: (1) countries with a high quality business environment and a positive domestic (immobile) tax base at hand (2) a country with a low quality business environment and no domestic tax base. The countries compete in tax rates to maximize tax revenues. In the economy, there exists a positive mass of mobile firms, who, in response to the set of announced tax rates, allocate a fixed level of production between the countries to maximize post-tax profits. The difference in productivity between high and low quality country is heterogeneous across firms, however, every firm prefers production in a high quality country on real economic grounds. Based on the allocation of production, a firm is allowed to shift profits as determined by the regulation in place. Broadly speaking, the regulation allows a firm that produces a share s in a country to claim up to a share $(1+x)s$ of its profits there. The model is solved as a subgame perfect Nash-equilibrium, conditional on the degree of the policy's limitation on profit shifting, the x.

In equilibrium, the low quality country always undercharges the competitors' tax rates due to its competitive disadvantage and the lack of a domestic tax base. Mobile firms with a sufficiently small productivity difference allocate parts of their production to the low quality country in order to reduce their tax bill. The effect of a lift in the profit shifting limitation on production efficiency can be broadly split up into three intuitive mechanisms. First, keeping the tax rates and firms' locations fixed, a lift in the limitation increases efficiency, because firms, who would have produced in the low quality country regardless, only need to allocate a smaller share of production there in order to retain the same ability to shift profits. Second, taking into account firms' location responses, a lift in the limitation decreases efficiency, because more firms are allocating parts of their production to the low quality country, incentivized by the higher shifting-per-production unit rate $(1+x)$. Third, taking into account countries' strategic responses in tax rates, a lift in the limitation decreases efficiency. High productive countries lose the ability to generate tax revenues from the mobile firms based on their competitive advantage and, in response, shift tax burden to the immobile tax base by increasing tax rates. Consequently, even more and less productive firms start to allocate parts of their production to the low quality country. The first two effects would still generate a monotonic decrease in the inefficiency over the whole support of intermediate policies. The third, strategic effect is only present for a sufficiently tight limit on profit shifting, because from a certain limitation onward, the high quality countries already selected out of competition for the mobile firms and shift the full tax burden to domestic firms. If present, the strategic effect is first-order and so, generates the non-monotonicity in production efficiency.

These theoretical results have direct policy implications. Assume a policymaker is interested in the containment of profit shifting and/or the maximization of global production efficiency. According to my results, there are only two candidates for an optimal regulation. Either deregulate profit shifting completely to achieve production efficiency, or set the regulation as tight as possible to contain profit shifting at a *locally maximized* efficiency.

In contrast, past years have seen the introduction of intermediate profit shifting policies, most prominently the modified Nexus approach. In a second set of results, this paper provides a novel rationale for the introduction of such policies. I ask the following Research Question: 'Who gains from the introduction of a limited degree of profit shifting?'. I will focus on the set of mobile firms.

In order to answer the question, I analyze the firms' marginal profits, starting from the benchmark of Proportional Apportionment. In general, a lift in the limitation has two counteracting effects on the firms' profits. On the one hand, firms gain from the enhanced flexibility in profit shifting. Given that firms operate under a higher shifting-per-production unit rate, profits increase: (1) either due to the reclaimed ability to avoid taxes by shifting to the low quality country (2) or due to the decrease of production in the low quality country, given a fixed amount of profits being shifted. On the other hand, firms lose from the increase in the economy's level of taxation. High quality countries charge higher tax rates to shift burden to their immobile tax base and low quality countries adjust tax rates upward by strategic complementarity.

The size of the overall marginal profits is mainly determined by the former effect and depends non-monotonically on the firms' productivity difference between high and low quality countries. In order to gain from a lift in the limitation, a firm needs to have: (i) a drop in productivity from allocating a unit of production to the low quality country (ii) a significant share of production in the low quality country. Firms with a small productivity difference fail on condition (i), firms with a large productivity difference on condition (ii). Concrete, a firm with a large productivity difference would gain significantly from allocating back parts of the production to the high quality country. However, these firms do not allocate capital to the low quality countries in the first place. Firms with a small productivity difference do so. But, for these firms allocating capital back does not really free up resources, they are more concerned by the upward trend in tax rates. There will be a subset of firms with intermediate productivity differences, who allocate a substantial share of production to the low quality country, but generate a significant productivity loss per production unit. My results show that it will be these firms, who gain the most from the lift in the limitation, while firms at the boundaries of the productivity difference support are certain to lose profits.

A collective action is more likely to be implemented, if gains from the action are concentrated on a small subgroup of agents, because concentration mitigates free-rider problems (Olson, 2009). A sizable strand of the literature on lobbying has applied these arguments to their field: Lobbying activities will be most likely successful, if gains from a policy are concentrated on a small subgroup of firms. Bringing these arguments together with the results of my paper, policies with limited profit shifting provide a breeding ground for lobbying activities[5]. In this light, this paper's theoretical results could provide a rationale for the introduction of policies

[5]The skepticism against limited profit shifting policies and the influence of business representatives has been raised previously

with limited profit shifting, in cases in which a standard policymaker's objective is incapable of.

The paper is located at the intersection of two different strands of the literature. On the one hand, the project adds to a sizable literature on the importance of general equilibrium effects in the evaluation of supplementary policies in the field of international taxation. As an example, Haufler and Runkel (2012) model the interaction of statutory tax rates and thin-capitalization rules in a tax competition framework[6]. The authors find that a coordinated tightening of thin capitalization rules makes the competition in statutory tax rates more aggressive, (partly) counteracting the positive budgetary impact of thin capitalization rules. My paper discusses the interaction of tax rates and limitation on profit shifting and shows that a lift in the limitation on profit shifting provides an adverse disincentive for high productive countries to compete for mobile tax revenues. On the other hand, the paper follows the literature on why certain countries do not follow the predicted 'Race to the bottom' in capital (corporate) tax rates. The main insights from this literature are neatly captured in Marceau et al. (2010). The authors analyze non-preferential tax competition for mobile and immobile capital. Countries differ in their endowment of the (immobile) domestic tax base. In equilibrium, capital tax rates are positively related to the size of the domestic tax base, reflecting the opportunity costs of international taxation in terms of foregone domestic tax revenues. In my model, the difference in the tax rate elasticity of domestic and mobile tax base is endogenous to the degree of profit shifting. The weaker the limitations are, the larger the difference, so that the opportunity costs from international taxation increase and countries optimally shift tax burden to their domestic tax base.

The remainder of the paper is structured as follows: Section 2 introduces the model in parts, Section 3 constructs the subgame-perfect Nash equilibrium. Based on the equilibrium, Section 4 discusses the main results of this paper, the effect of a change in the limitation on profit shifting on production efficiency and marginal profits of firms. Furthermore, Section 5 complements these results by the discussion of the effect on other relevant statistics, e.g. tax revenues of high and low productive countries. Section 6 goes in for sensitivity of the results and lines out possible extensions. Finally, Section 7 concludes.

2 Model: Tax Competition

A continuum of firms allocates capital across their worldwide production facilities to develop a project with a fixed level of taxable profits. The allocation of taxable profits across countries is regulated by a production-based apportionment regime, potentially allowing for a limited degree of profit shifting. Countries offer investment opportunities for the firms and charge corporate taxes. The model at hand is designed to determine the impact of the profit shifting limitation on the economy, incorporating the direct effects through changes in firms' behavior as well as the strategic effects through changes in the statics of the tax competition game. The following section introduces the parts in detail and lays out the model, the final part discusses the main assumptions of the model.

in the discussion on the modified Nexus-approach, see BEPS Monitoring Group (2015).

[6] Other examples are Haufler et al. (2018) on the indirect impact of controlled-foreign-company rules or Brekke et al. (2017) on the indirect impact of rules regarding the deductibility of equity costs.

2.1 Timing

The model spreads out over three stages. The set-up leans on the the design of the OECD modified Nexus-approach, however, for the ease of exposition it differs from the actual policy on certain dimensions. The differences and their implications are discussed in Section 6. *Stage 1:* Two homogeneous countries, A and B, simultaneously choose their corporate tax rates to maximize tax revenues. *Stage 2:* A single country, C, optimally chooses its corporate tax rate to maximize tax revenues. C potentially differs from A and B in terms of size and productivity. Finally, C fully observes the decisions on previous stages. *Stage 3:* The firms in the economy observe the set of policies and determine their investment profile and profit shifting activities to maximize post-tax profits.

2.2 The firms - production

A unit mass of firms in the economy decides about whether and how to invest in a project with fixed taxable profits Π. In order to develop the project, a firm needs to generate an output weakly larger F. Firms generate output by investing capital across the three countries in the economy, {A;B;C}. In particular, for a firm i, investing K_A units of capital in A translates into $\theta^i_A K_A$ units of output. The per-unit rental rate of capital is symmetric and given by r. Putting the parts together, for a firm i, the project is developed under the constraint:

$$\sum_{j \in \{A;B;C\}} \theta^i_j K_j \geq F \tag{1}$$

The double-indexing of θ indicates that the productivity of capital is firm-country specific. Initially, firms are split up in mobile vis-à-vis domestic firms. Throughout the model, the mass of firms domestic to C is set to zero. In difference, A and B each host a positive mass d of domestic firms. Domestic firms have a zero productivity of investment outside the home country and home productivity is normalized to $\overline{\theta} > 0$. The remaining mass of firms, $m = 1 - 2d$, is mobile between the countries in the economy. The productivity of mobile firms in A or B is symmetric and equal to the 'domestic firm-home country' productivity $\overline{\theta}$. In difference, productivity in C is heterogeneous across mobile firms and captures their idiosyncratic degree of mobility. In particular, the mass of mobile firms is uniformly distributed on an interval $[0, 1]$. The position of an individual firm on the interval is depicted by $f \in [0, 1]$. From the position on the interval, the productivity in C derives as $\theta^f_C = \frac{\overline{\theta}}{1+f\overline{\theta}\Delta}$; the smaller f, the higher the productivity in C. Thus, the degree of a firm's mobility decreases in f. For the rest of the model, I will keep indexing the mobile firms with respect to their position on the unit interval, f. Finally, the overall degree of productivity differentiation in the economy is governed by Δ; for $\Delta \to 0$, the productivity of mobile firms is symmetric across countries and firms.

2.3 The firms - profit shifting

Profit shifting describes the effort of mobile firms to reduce the global tax bill, given an allocation of production. The reduction in tax payments is achieved by claiming an as high as possible share of profits in low tax jurisdictions. In fact, firms choose the share of profits claimed in the respective countries, given by $\{A_A; A_B; A_C\}$, to minimize the average tax rate $t_{av} = A_A t_A + A_B t_B + A_C t_C$. The country-specific shares are a choice of the firms. The bounds on the choice set are determined by the allocation of production and the profit shifting regulation into place. In particular, the choice sets are determined by two constraints.

First, the country-specific share A_j has to be chosen from $[0, \overline{A_j}]$, where the upper bound is given by:

$$\overline{A_j} = \min\left\{\alpha\frac{\theta_j^f K_j}{F}; 1\right\}$$

The upper bound depends on the share of production in j, $\frac{\theta_j^f K_j}{F}$, and an *uplift parameter*, α. This parameter allows firms with a positive investment stock to move profits in the tax jurisdiction of j, dis-proportional to the actual share of production. So, the uplift controls the limitation on profit shifting. The higher the uplift, the more profit shifting is tolerated in the economy. Finally, a firm is never allowed to claim more than full profits in a country, which gives the second part of the upper bound's max-operator. The formulation of the upper bound is in line with the OECD's modified Nexus.

Second, the country-specific shares have to add up to 1, $A_A + A_B + A_C = 1$. The sum is not allowed to be smaller than 1, there is no tax evasion. The sum is not allowed to be larger than 1, there is no double taxation.

Consider a simple example to understand how profit shifting works in the model and how it is shaped by the regulation. Assume firm f has evenly split production across the three countries, $\theta_A^f K_A = \theta_B^f K_B = \theta_C^f K_C = \frac{F}{3}$. Further assume a sequence of tax rates so that $t_A > t_B > t_C$. Table 1 presents sets of optimally chosen shares, given the allocation of production and different uplift parameters.

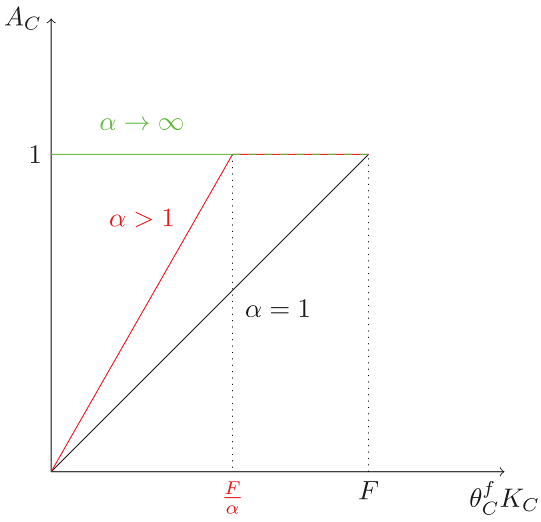
	A_A	A_B	A_C
$\alpha = 1$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$\alpha = 2$	0	$\frac{1}{3}$	$\frac{2}{3}$
$\alpha = 3$	0	0	1
$\alpha = 4$	0	0	1

Table 1: Choice of shares

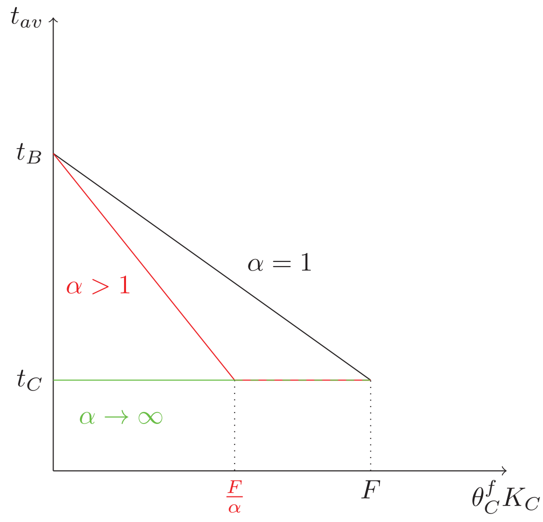
The table shows the optimal profit shares for different levels of α, given an equal split of production and $t_A > t_B > t_C$.

Given its allocation of production, a firm claims as much profits as possible in the country with the lowest tax rate. The larger α, the more freedom a firm has in its choice to move profits. However, at some level of α, in the example $\alpha = 3$, the limitations on profit shifting are sufficiently weak so that, given its allocation of production, the firm is allowed to claim full profits in the country with the lowest tax rate. From that point on further increases in α do not change the choice set and so, the optimal choice of shares.

In difference, consider changes in the allocation of production, but fix the α. Further, keep the sequence of tax rates from the example above. Figure 1a plots the optimal share A_C as a function of the production level. Figure 1b visualizes the resulting average tax rate, assuming that the firm splits production between B and C only.



(a) Optimal share



(b) Average tax rate

Figure 1: Profit allocation

If a firm produces $\frac{F}{\alpha}$ in a certain country, it is allowed to claim full profits there, any further increase in production will not affect the upper bound of the share's choice set. This threshold will become important in the later analysis and therefore, should be kept in mind.

With regard to the benchmark regulations, see that a profit shifting regulation with $\alpha = 1$ implies a classic Proportional Apportionment. On the other side of the spectrum, for $\alpha \to \infty$ the regulation approaches a situation with unregulated profit shifting. Accordingly, the analysis of this paper will focus on regulations with $\alpha \in [1, \infty)$.

2.4 The countries

In the model, countries potentially differ on three dimensions: size of the domestic tax base, productivity of mobile firm investment, stage of appearance. A and B are homogeneous in each of the three dimensions. Both have a domestic tax base d, mobile firms have a symmetric productivity of investment $\overline{\theta}$ and finally, both countries simultaneously choose tax rates on stage 2 of the game. In difference, C is distinct from the other countries in all the three dimensions. First, the domestic tax base of C is set to zero; C has a negligible domestic sector of production. Second, mobile firms have a weakly lower productivity of investing in C; C has a competitive disadvantage in attracting mobile investment on real economic grounds. Third, C moves after the other two countries; C is more flexible in policy-making. The different stages of appearance are kept throughout the model. The difference in the size of the domestic tax base is parameterized by d, in particular for $d \to 0$, the model approaches a situation of symmetry in size among the countries. Similarly, for $\Delta \to 0$, differences in the productivity of investment vanish. Throughout the model, I restrict the parameter space such that $d \leq m$ and $\Delta \leq 0.5\frac{\Pi}{rF}\left(1 - \frac{rF}{\overline{\theta}\Pi}\right) \equiv \Delta^{max}$. The restriction on d is to generate interior solutions and without loss of generality. The choice of Δ^{max} is made to define mobility of a firm and will be explained in detail later in the paper.

2.5 Discussion

To keep the model tractable, I impose a number of simplifying assumption. I use this section to discuss the main assumptions. Section 6 will elaborate on possible extensions of the model.

Production: The model considers the firms' investment decision as a cost-minimization problem. For a given level of output, a firm allocates capital internationally to minimize tax and capital costs. Modelling profits in discrete unit, e.g. profits from a project, is widespread in the literature on the development of intangible assets. In particular, it captures the idea of intangibles being a fixed factor of production. I follow this approach, because the empirical literature on profit shifting has shown that profits from intangible assets are the most relevant source of tax planning activities. In line, this model's assumption on the production process could be interpreted as a standard Poisson process, with arrival rate $\lambda = 1$ for an output larger F and 0 otherwise. This process could be formulated more general, e.g. by modeling λ as a continuous function of the research output. None of the qualitative results obtained in this paper is sensitive to these simplifying assumptions. Finally, the assumptions are in line with recent approaches in the theoretical literature on international taxation of revenues from intangibles, e.g. Juranek et al. (2018).

Competitive sequence: The theoretical literature on tax competition predominantly works under the assumption of simultaneous tax setting. However, motivated by empirical work like Altshuler and Goodspeed (2015) there exists a sizable body of literature on the impact of sequential choices on tax competition, as summarized in Keen and Konrad (2013). The model at hand assumes Stackelberg-leadership of the larger and more productive countries, motivated e.g. by the recent evidence on the Luxembourgian tax rulings, Karnitschnig and Van Daalen (2014). Flipping the competitive sequence would not change any of the main qualitative results. Imposing simultaneous decision-making of all countries would significantly complicate the analysis, however, does not seem to change the main qualitative results as well. Competition on the leader stage is formulated as a duopoly, competition on the follower stage as a monopoly. This modelling choice allows to capture some intuitive mechanisms in a mathematically tractable way. The main qualitative results hold with different within-stage competition designs.

3 Equilibrium analysis

The following section derives the subgame-perfect Nash equilibrium of the tax competition game defined by:

Definition 1. *A subgame-perfect equilibrium is a profile of optimal tax rates for high productive countries,* t_H^**, best responses in tax rates of the low productive country,* $t_L^*(t_H)$*, firms' best responses in capital allocations* $K^*(t_H, t_L)$ *and profit allocations* $A^*(t_H, t_L)$*, such that:*

1. *For every high productive country* $h \in \{\text{A;B}\}$*, the tax rate* t_h^* *maximizes tax revenues given* t_{-h}^* *and the best responses of the low productive country and of the firms.*

2. *For the low productive country C, the best response tax rate* $t_L^*(t_H)$ *maximizes tax revenues of C given the best responses of the firms.*

3. *For every firm i, the best response capital allocation $K_i^*(t_H, t_L)$ and best response profit allocation $A_i^*(t_H, t_L)$ maximize post-tax profits.*

I will solve the model by backwards induction and delegate all the proofs to the Appendix.

3.1 The firms

Conditional on participation, the best response of firm i is given by the capital allocation $K_i^*(t_H, t_L)$ and the profit allocation $A_i^*(t_H, t_L)$ that solve:

$$\max_{K_i, A_i} (1 - t_{av}) \Pi - r \sum_{j \in \{A;B;C\}} K_j \tag{2}$$

$$\text{s.t.} \quad (\lambda): \sum_{j \in \{A;B;C\}} \theta_j^i K_j \geq F \qquad (\mu): \ A_j \in [0, \overline{A}_j] \qquad (\kappa): \sum_j A_j = 1$$

Where t_{av} and $\overline{A}_j$ were defined in Section 2.3. Participation profits of a domestic firm[7] are given by $(1-t_A)\Pi - r\frac{F}{\overline{\theta}}$, since, by definition of domesticity, the optimal capital allocation to enter the market is $K_d^*(t_H, t_L) = \left\{\frac{F}{\overline{\theta}}; 0; 0\right\}$, with an according allocation of profits $A_d^*(t_H, t_L) = \{1; 0; 0\}$. Imposing an outside profit of 0, a domestic firm enters the market, iff $t_A \leq 1 - \frac{r}{\overline{\theta}} \frac{F}{\Pi} \equiv \bar{t}$. Throughout the model, I will label $\bar{t}$ as the *domestic monopoly rate* and, without loss of generality, focus on $\{t_A, t_B, t_C\} \in \left[0, \bar{t}\right]^3$.

Optimal profit shifting of the mobile firm is in line with the observations of Section 2.3. Given an international allocation of production, a firm allocates profits to minimize the average tax rate. Technically this means that for a sequence of tax rates $t_j > t_k > t_l$, the optimal choice of profit allocation is given by: $A_j^* = \overline{A}_j$; $A_k^* = \min\{\overline{A}_k, 1 - A_j^*\}$; $A_l^* = \min\{\overline{A}_l, 1 - A_j^* - A_k^*\}$. From now on, we discuss the optimal capital allocation, taking into account an optimal allocation of profits with regard to the previous decision rule.

In the choice of the capital allocation, a mobile firm trades off the marginal impact on the average tax rate, $-\frac{\partial t_{av}}{\partial K_j}\Pi$, against the real economic impact of an investment, $-r + \theta_j^f \lambda$. In particular, a mobile firm f invests in j, iff:

$$\left(\frac{\partial t_{av}}{\partial K_{-j}} - \frac{\partial t_{av}}{\partial K_j}\right) \Pi \geq \left(\theta_{-j}^f - \theta_j^f\right) \lambda \quad \forall j \in \{A; B; C\} \tag{3}$$

The marginal average tax rate has a discontinuity at $K_j = \frac{F}{\theta_j^f \alpha}$. At the point of discontinuity, the firm is just allowed to claim full profits in j. For investment levels above, the marginal average tax rate is 0. For investment levels below, the marginal average tax rate is constant and has the same sign as $(t_j - t_{-j})$. These insights are used in the formulation of the first lemma:

Lemma 1. capital allocation - intensive margin
If a mobile firm f invests in C (the low productive country), it will invest as much that it is just allowed to shift full profits, $K_f^*(C) \in \left\{0; \frac{F}{\alpha \theta_C^f}\right\}$.

Given that a unit of mobile capital is weakly less productive in C, a mobile firm will only invest in C due to the impact on the average tax rate. Given linear production and profit shifting regulation, once the firm starts investing in C, it will extend the investment stock as long as all profits are allowed to be claimed in C. Further, applying (3) to the specific countries, two conditions for investment emerge:

[7] Without loss of generality, I will consider a firm domestic to A.

Lemma 2. capital allocation - extensive margin

(i) *A positive mass of mobile firms pays taxes in A, only if* $t_A \leq t_B$.

(ii) *A positive mass of mobile firms invests and pays taxes in C, if and only if* $t_C < \min\{t_A; t_B\}$.

For the class of mobile firms, A and B offer symmetric investment opportunities, such that the decision to invest and pay taxes in either of them two is fully due to the tax incentives. Imposing a tie-breaking rule that for equal tax rates, firms prefer moving to A, gives the intuition of (i). On the other hand, for an investment in C, firms have to be compensated for the loss in investment productivity. This compensation has to be done by the tax bill, which gives the intuition for (ii), where sufficiency is due to the symmetric productivity of the perfectly mobile firm. Building on the set of Lemmas, the first Proposition of this paper is formulated as follows[8]:

Proposition 1. capital allocation - best response
For any set of tax rates in high productive countries $t_H = \{t_A; t_B\} \in [0, \bar{t}] \times [t_A, \bar{t}]$

- *for* $t_C \geq t_A$, *the symmetric best responses in capital and profit allocation for all mobile firms are given by:*

$$K_f^* = \left\{K_f^*(A); K_f^*(B); K_f^*(C)\right\} = \left\{\frac{F}{\overline{\theta}}; 0; 0\right\} \qquad A_f^* = \left\{A_f^*(A); A_f^*(B); A_f^*(C)\right\} = \{1; 0; 0\} \qquad \forall f \in [0,1]$$

- *for* $t_C < t_A$, *there exists a mobile firm with* $f^{thr} = \min\left\{1; \frac{\alpha}{\Delta}(t_A - t_C)\frac{\Pi}{rF}\right\}$, *such that the best response capital and profit allocation are given by:*

$$K_f^* = \left\{\left(1 - \frac{1}{\alpha}\right)\frac{F}{\overline{\theta}}, 0, \frac{1}{\alpha}\frac{F}{\theta_C^f}\right\} \qquad A_f^* = \{0; 0; 1\} \qquad \forall f \in [0, f^{thr})$$

$$K_f^* = \left\{\frac{F}{\overline{\theta}}; 0; 0\right\} \qquad A_f^* = \{1; 0; 0\} \qquad \forall f \in [f^{thr}, 1]$$

The first part of Proposition 1 follows directly from joining the two statements of Lemma 2. Regarding the second part, Lemma 1 has shown that in case of an investment in C, firms always invest a share of production so that they are just allowed to shift full profits. Since a firm needs more capital to produce a certain level of output in C, the global capital demand of firm f is higher when investing in C. In expression, the additional capital costs from investing in C are given by:

$$ACC^f = r\left(\left(1 - \frac{1}{\alpha}\right)\frac{F}{\overline{\theta}} + \frac{1}{\alpha}\frac{F}{\theta_C^f}\right) - r\frac{F}{\overline{\theta}} = \frac{1}{\alpha}\left(\frac{1}{\theta_C^f} - \frac{1}{\overline{\theta}}\right) rF = f\frac{\Delta}{\alpha} rF \tag{4}$$

The additional capital costs increase in f. For the marginal firm f^{thr}, the additional capital costs match the associated reduction in the tax bill, $(t_A - t_C)\,\Pi$, and so, firms with a larger f centralize the investment in A. In addition, the model is flexible enough to capture situations in which all the firms generate strictly positive marginal profits from shifting profits to C, $\frac{\alpha}{\Delta}(t_A - t_C)\frac{\Pi}{rF} > 1$.

In order to understand the impact of different profit shifting regulations on behavior and profits of the mobile firms, consider two uplift levels, α_{low} and α_{up}, with $\alpha_{up} > \alpha_{low}$. Using the functional form provided in Proposition 1, it holds that $f^{thr}(\alpha_{up}) > f^{thr}(\alpha_{low})$; as the share of investment in C required for profit shifting, $\frac{1}{\alpha}$, decreases, more firms invest in C. Assuming that $f^{thr}(\alpha_{up}) < 1$, the continuum of mobile firms can be split up in three parts. First: Firms on $[0, f^{thr}(\alpha_{low})]$ invest in C under a tighter profit shifting regulation already,

[8] For the ease of exposition and without loss of generality, the proposition is formulated in terms of A.

labeled as *inframarginal firms*. Second: Firms on $\left[f^{thr}(\alpha_{low}), f^{thr}(\alpha_{up})\right]$ invest in C under the alleviated profit shifting regulation only, labeled as *marginal firms*. Third: Firms on $\left[f^{thr}(\alpha_{up}), 1\right]$ do not invest in C even under the alleviated profit shifting regulation, labeled as *sticky firms*.

If a policy lifts the limitation on profit shifting, the inframarginal firms will fetch back capital to their more productive facilities, retaining the same profit shifting allowance. This adjustment of the capital allocation on the *intensive margin* decreases the additional capital costs in proportion to their initial level. Since for inframarginal firms, initial capital costs increase in f, the marginal profits from a lift in the limitation increases in f within the class of inframarginal firms. In difference, marginal firms start investing in C after the lift in the limitation, they adjust their investment on the *extensive margin*. So, their gain from the uplift is the reclaimed ability to use the low productive-low tax country as a vehicle to cut the tax bill. However, capital costs increase in f and thus, the marginal profits from a lift in the limitation decrease in f within the class of inframarginal firms. Finally, sticky firms neither react on the extensive nor on the intensive margin. Thus, profits are not directly affected by the lift in limitation. Figure 2 graphically presents the additional capital costs and tax gains from investing in the low productive country. Further, it depicts the situation of different limitations on profit shifting and their impact on the firms' post-tax profits.

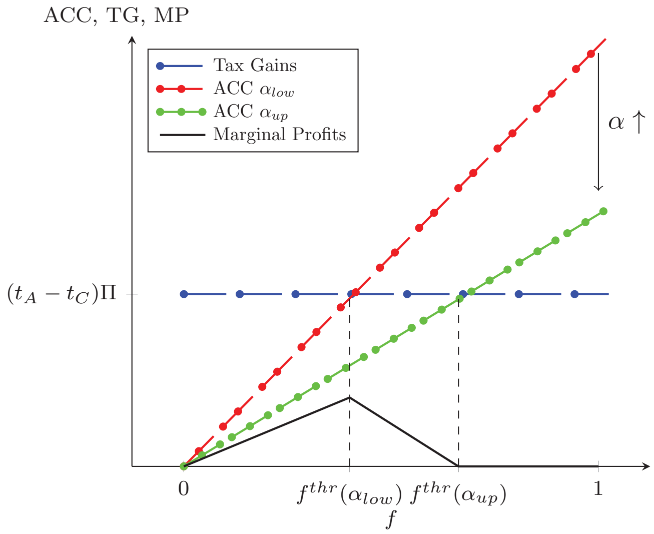


Figure 2: Additional Capital Costs, Tax Gains and Marginal Profits

The analysis has shown that mobile firms are affected by changes in the limitation on profit shifting in a non-monotonic way, depending on a firm's degree of mobility. Throughout the paper, I will keep using the distinction between inframarginal, marginal and sticky firms.

3.2 The low productive country

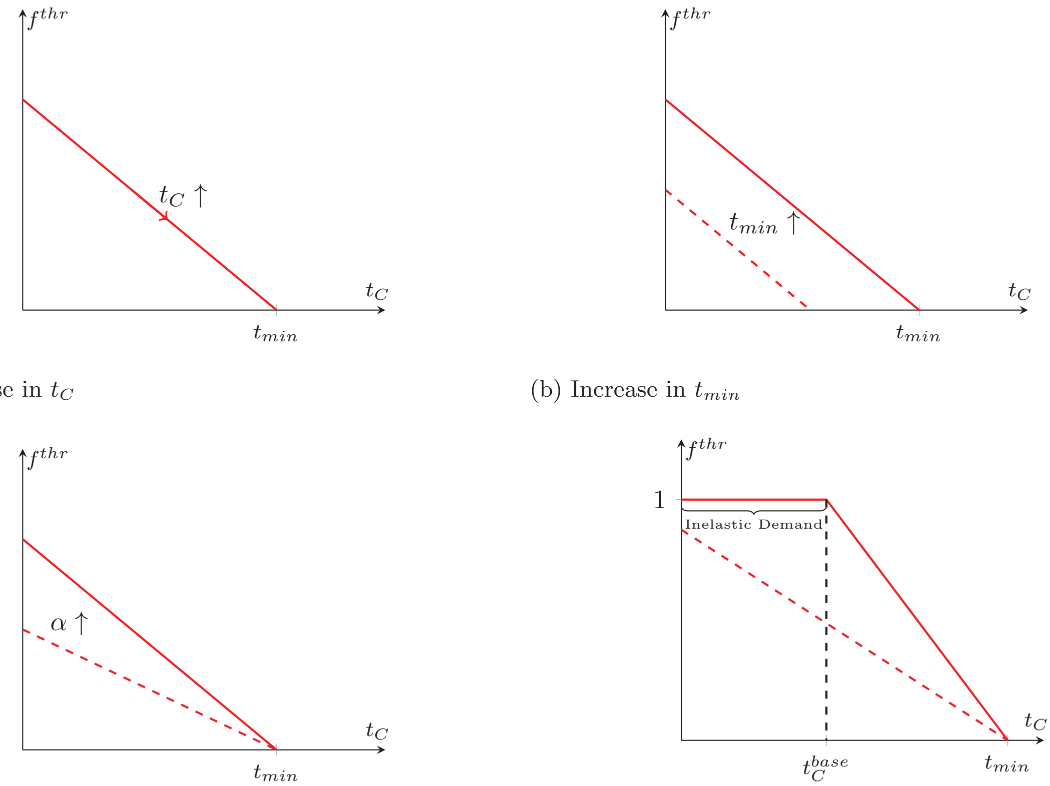
The best response of the low productive country C is the corporate tax rate $t_C^*(t_H)$ that solves:

$$\max_{t_C} TR_C = \begin{cases} 0 & \text{if} \quad t_C \geq \min\{t_A; t_B\} \\ t_C f^{thr}(t_H, t_C)\, m\Pi & \text{if} \quad t_C < \min\{t_A; t_B\} \end{cases} \tag{5}$$

Given that $f^{thr} > 0$, choosing $t_C \geq \min\{t_A; t_B\}$ is strictly dominated by any $t_C \in (0, \min\{t_A; t_B\})$[9]. Thus, during this section, the focus is on the lower part of (5).

The mass of firms shifting full profits to C is described by $f^{thr}m$. Regarding tax revenues, the mass, multiplied by the symmetric per-firm profits Π, describes the demand for profits claimed in C. In addition to a set of parameters, the demand depends on the tax rate of C as well as on the minimum tax rate of the high productive countries.

The demand is downward sloped in the tax rate of C. Ceteris paribus, the higher the tax rate, the fewer firms are willing to incur additional capital costs in order to enjoy the lower tax rate in C, as depicted in Figure 3a. The demand shifts upward in t_{min}. An increase in the minimum tax rate of high productive countries generates a larger cut in the tax bill from shifting to C and thus, increases the willingness to invest in C. The tax cut is symmetric across firms and so, the demand curve is shifted up parallel, as depicted in Figure 3b. Further, the demand rotates outward as α increases, keeping the choke price (tax rate) fixed. For t_C close to the choke price, only the most productive firms move to C. As discussed in Section 3.1, these firms gain the least from a lift in the profit shifting limitation and thus, the increase in the demand slows down the smaller the tax differential $(t_{min} - t_C)$, as depicted in Figure 3c. The maximum demand is exogenously capped at $f^{thr} = 1$. Thus, the structure of the demand potentially inhibits a kink. To the left of the kink, all mobile firms generate positive marginal profits from investing in C and thus, the demand for profits is locally inelastic. To the right of the kink, there exists a firm indifferent to an investment in C, which reinforces the properties described above, the situation of a kinked demand is represented in Figure 3d.



(a) Increase in t_C

(b) Increase in t_{min}

(c) Increase in α

(d) Kinked demand

Figure 3: The demand for profits in C

[9]From here on, for the ease of exposition, label $\min\{t_A; t_B\} \equiv t_{min}$

The following Proposition summarizes the main results of this subsection:

Proposition 2. low productive country - best response
For any set of tax rates in high productive countries $t_H = \{t_A; t_B\} \in [0, \bar{t}]^2$ *the low productive country competes for mobile firms with best responses in tax rates given by:*

$$t_C^* = \begin{cases} 0.5t_{min} & \text{if } t_{min} \leq 2\frac{rF}{\Pi}\frac{\Delta}{\alpha} \\ t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha} & \text{else} \end{cases}$$

In general, tax rates are strategic complements in the model; the best response tax rate of the low productive country strictly increases in the minimum tax rate of the high productive countries on the whole tax rate support. However, the level and the slope of the best response depend on the level of t_{min}. In the discussion of Figure (3d), it has been shown that the demand for profits potentially inhibits a kink at $t_C^{base} = t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha}$. Choosing a tax rate below t_C^{base} is strictly dominated by t_C^{base} given the locally inelastic demand. On the other hand, C chooses a tax rate on the downward-sloping part of the demand curve, only if the positive tax rate effect outweighs the negative tax base effect. Given the linearity of the demand, tax rate and base effect decrease in t_C and thus, C increases the tax rate above t_C^{base}, if and only if marginal tax revenues at the kink are positive, algebraically:

$$\underbrace{mf^{thr}\left(t_{min}, t_C^{base}\right)\Pi}_{\text{tax rate effect}} + \underbrace{t_C m \frac{\partial f^{thr}}{\partial t_C}\bigg|_{t_C = t_C^{base}}\Pi}_{\text{tax base effect}} \geq 0 \tag{6}$$

Using the functional form of f^{thr} it can be shown that the condition holds for $t_{min} \leq 2\frac{rF}{\Pi}\frac{\Delta}{\alpha}$ with a best response on the downward-sloping part given by $t_C = 0.5t_{min}$. For higher t_{min}, the condition fails and thus, C chooses t_C^{base} as the best response. For intuition, see that (6) is equivalent to $1 + \epsilon_{f^{thr}, t_C}|_{t_C = t_C^{base}} \geq 0$. Thus, the low productive country chooses to host the full mass of mobile firms, if the demand is relatively elastic on the downward-sloping part. Instead, for sufficiently inelastic demand functions, the low productive country decides to increase its charges and caters only a fraction of the mobile firms. A higher t_{min} increases the elasticity in absolute terms by allowing the low productive country to extract a larger profit margin per-firm. Thus, the tax base effect from hosting only a fraction of the mobile firms decreases and the low productive country abstains from becoming a niche provider with even higher margins.

The analysis has shown that, given its negligible domestic tax base, the low productive country always chooses its policies to attract mobile firms. Depending on the tax rates of high productive countries, the low productive country either locates in a *niche* hosting only a fraction of mobile firms with best response $t^n = 0.5t_{min}$ or serves the full *mass* of mobile firms with best response $t^m = t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha}$. I will keep using the terms mass and niche provider for the analysis in the remaining parts of the paper.

3.3 The high productive countries

For a high productive country h, the optimal corporate tax rate t_h^* solves:

$$\max_{t_h} TR_h = \begin{cases} t_h \quad d & \Pi \quad \text{if} \quad t_h > t_{-h} \\ t_h \left(d + m(1 - f^{thr}(t_h)\right) & \Pi \quad \text{else} \end{cases} \tag{7}$$

If a high productive country announces a tax rate higher than its direct competitor on Stage 1, it will generate revenues from the taxation of its domestic tax base only. If a high productive country underscores the competitor, it will additionally host and tax a mass $\left(1 - f^{thr}(t_h)\right) m$ of mobile firms. The mass of mobile firms, multiplied by the per-unit profits Π, constitutes the mobile demand for profits claimed in the high productive country's tax jurisdiction. The high productive country incorporates the best responses of the low productive country and thus, the mobile demand is expressed as a function of the policies of the high productive country only. If the low productive country chooses to act as a niche provider, an increase in t_h reduces the mobile demand for profits claimed in h, given that the low productive country increases its tax rate by less than unity to balance out the elasticity. If the low productive country acts as a mass provider, h does not generate mobile tax revenues. It has been shown that due to the increase in the per-firm profit margin, the low productive country acts to host the full mass for a sufficiently high $t_h > t_h^{thr}$. Thus, the demand for profits in a high productive country is downward-sloping until t_h^{thr} and flattens out at level d thereafter, as depicted in Figure 4. Regarding the profit shifting regulation, the demand rotates inward for an increase in α, keeping fixed the demand at $t_h = 0$, as depicted in Figure 4. On the downward-sloping part, the reduction increases in t_h, given that the marginal firm's investment decision is more sensitive to changes in the limitation on profit shifting for higher tax rates.

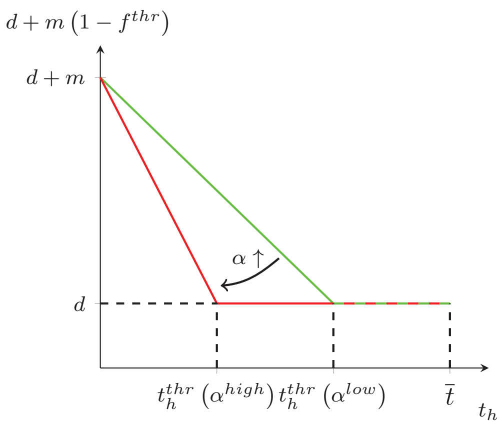


Figure 4: The demand in high productive countries

The upper bound on the degree of differentiation, Δ^{max}, guarantees that $t_h^{thr}(\alpha = 1) \leq \bar{t}$; a mobile firm never pays the domestic monopoly rate in a high productive country. So, by reverse engineering, the restriction on the parameter space gives the definition of mobility in the model. The following Proposition captures the main results of this subsection:

Proposition 3. tax rates - equilibrium
There exists a level of limitation implied by $\alpha = 0.5\frac{1}{d}\left(1 + \frac{d^2}{m}\right)\Delta\left(\frac{1}{\theta} + \Delta\right)\frac{\frac{r}{\theta}F}{\Pi - \frac{r}{\theta}F} \equiv \alpha^{thr}$*, such that:*

- *for* $\alpha \geq \alpha^{thr}$ *there exists a unique pure strategy equilibrium in which high productive countries do not compete for mobile firms, instead charge the domestic monopoly rate* $\bar{t}$*.*
- *for* $\alpha < \alpha^{thr}$ *there exists an equilibrium in mixed strategies with equilibrium support* $\Omega_A = \Omega_B = \left\{[t', t'']; \bar{t}\right\}$*, where* $t'' < \bar{t}$*. In the mixed strategy equilibrium, high productive countries compete for mobile firms with a strictly positive probability* $\left(1 - Pr\left(\bar{t}\right)\right)$ *and choose the tax rate from a convex interval* $[t', t'']$ *under a cumulative density* $G(t)$ *with* $G(t') = 0$ *and* $G(t'') = 1 - Pr(\bar{t})$*. With remaining probability* $Pr\left(\bar{t}\right)$ *high productive countries do not compete for mobile firms, instead charge the domestic monopoly rate* $\bar{t}$*.*

Pure-strategy equilibrium: As depicted in Figure (4), to generate mobile tax revenues, a high productive country has to decrease its domestic monopoly rate by a discount weakly larger than $(\bar{t}-t_h^{thr})$. The discount is necessary to attenuate the profit margin of the low productive country, pushing her to the niche. At α^{thr} the negative domestic tax base effect, triggered by the discount, matches the mobile tax revenues. Consequently, for higher uplift levels a high productive country maximizes tax revenues by charging the domestic monopoly rate from its domestic tax base, irrespective of the other high productive country's tax rate. The threshold relies on three interpretable terms. The first term captures the relative importance of the domestic tax base. The term decreases in d, reflecting the fact that there are smaller incentives to provide a discount, if the domestic tax base gains relative importance. The second term captures the degree of differentiation across firms. For a higher differentiation, mass is shifted to firms with a low degree of mobility. Given that less mobile firms are more sticky in high productive countries, the required discount at every uplift level decreases. Therefore, a high productive country is willing to accept a larger α resulting in the same overall discount. The third term captures the relative importance of the production costs for any mobile firm. The higher the ratio of production costs to revenues, the more the optimal investment pattern of any firm aligns with its relative productivity and thus, the lower the necessary discount at every level of uplift. Consequently, the more important investment costs are for the firms, the larger the α for which a high productive country is willing to push the follower to the niche[10].

Tax Revenues in the mixed strategy-equilibrium: For uplift levels below the threshold, there exists a set of tax rates $[t', t''] \equiv \mathbb{T} \subset (0, t_h^{thr})$ such that mobile tax revenues weakly exceed the negative domestic tax base effect. There is no pure strategy-equilibrium with $t \in \mathbb{T}$. A high productive country always has the incentive to marginally undercut the other high productive country's tax rate to earn the mobile tax revenues. Further, given the positive mass of domestic firms and the imposed tie-breaking rule, deviating from a symmetric equilibrium in $t = t'$ to the domestic monopoly rate increases tax revenues for B. However, there exists a mixed strategy-equilibrium. Using (7), for any t in the equilibrium support, expected revenues are given by:

$$G(t)td\Pi + (1 - G(t))\, t \left(d + m \left(1 - f^{thr}(t)\right)\right) \Pi \tag{8}$$

Here, the (symmetric) cumulative distribution function $G(t)$ represents the probability that a high productive country with tax rate t is underscored by its direct competitor. Proposition 3 states that the domestic monopoly rate is part of every equilibrium support. It has been shown that $f^{thr}\left(\bar{t}\right) = 1$ unambiguously holds and thus, when charging $\bar{t}$ a high productive country earns the domestic monopoly tax revenues, irrespective of the competitor's tax rate. By the construction of a mixed strategy equilibrium, expected revenues have to be the same at any t in the equilibrium support. Thus, expected tax revenues from the mixed strategy equilibrium are equal to the domestic monopoly tax revenues.

Impact of α on the mixed strategy equilibrium: The equilibrium strategies of high productive countries react to the level of α. In particular, at any tax rate in the equilibrium support, the mixing probabilities have to adjust accordingly to preserve (8). Algebraically this requires:

$$\frac{\partial G(t)}{\partial \alpha} = -\frac{1 - G(t)}{1 - f^{thr}} \frac{\partial f^{thr}}{\partial \alpha} < 0$$

[10]See that there exist tuples of Δ and d, such that $\alpha^{thr} < 1$. In the remainder of the paper, I will focus on parameter tuples such that $\alpha^{thr} \geq 1$. In particular, the parameter space is chosen such that $\forall \Delta \in (0, \Delta^{max})\ \exists \mathrm{d}^* > 0, \text{s.t.}\ \forall \mathrm{d} \leq \mathrm{d}^*,\ \alpha^{thr} \geq 1$.

Further, given that $\Pr(\bar{t}) = (1 - G(t''))$, it holds that $\frac{\partial \Pr(\bar{t})}{\partial \alpha} > 0$. So, technically, an increase in α triggers a First-Order Stochastic Dominant shift of the mixing probability. Intuitively, the probability of hosting the mobile tax base has to increase in order to compensate for the loss in the actual inflow of profits. For the global equilibrium of the model, this shift has two distinct implications. First, the probability of the low productive country hosting the full mass of mobile firms, $\Pr(\bar{t})^2$, increases in α. Second, conditional on having a high productive country competing in the economy, the expected tax rates increase. If the uplift level approaches the threshold, the mixed strategy-equilibrium converges to the pure strategy-equilibrium, i.e. $\lim_{\alpha \to \alpha^{thr}} G(t) = 0 \;\; \forall t < \bar{t}$.

Relation to the literature: The competition between high productive countries is modelled as a price competition with loyal customers, in the spirit of Varian (1980). Such models have been previously used in the literature on international taxation (Elsayyad and Konrad, 2012). In order to evaluate the impact of limited profit shifting, my model extends previous work by the inclusion of the low productive country. In consequence, the size of the price-sensitive segment in the competition between high productive countries is not fixed anymore, but depends on the level of the tax rate in high productive countries and the implemented regulation on profit shifting. The modification is necessary to evaluate the impact of limited profit shifting. The construction of the tax competition equilibrium in parts still follows the work of Narasimhan (1988), however, their approach is adjusted multiple times to take into account this model's extension.

In sum, this subsection has derived the equilibrium policies of high productive countries, taking into account the best responses down the competitive sequence. The construction of a mixed strategy equilibrium follows common practice in Bertrand-type models with loyal customers, however, was modified for the sequence of the model. In a mixed strategy equilibrium, a high productive country competes for mobile profits with probability $\left(1 - \Pr\left(\bar{t}\right)\right)$. Instead, in a pure strategy equilibrium, a high productive country never competes for mobile profits and charges the domestic monopoly rate throughout.

4 The effect on main economic outcomes

This section discusses the impact of the limit on profit shifting on two main economic outcomes, efficiency of the global capital allocation and post-tax profits of mobile firms. The analysis starts to focus on these two statistics as they potentially provide a rationale for the introduction of policies with a limited amount of profit shifting. Increasing production efficiency is the official rationale of current policies with limited profit shifting, like the OECD's modified nexus. However, the results of the following analysis will show that the impact of limited profit shifting policies on production efficiency is ambiguous. In particular, for a degree of limitation like it was imposed by the OECD, my model predicts a decrease in efficiency compared to a classic Proportional Apportionment. The second part of this section provides a different rationale for the emergence of policies with limited profit shifting. This rationale will depend on the distribution of marginal profits across firms of different mobility.

4.1 Production Efficiency

This subsection analyses how the limitation on profit shifting affects production efficiency. Given that the idea behind policies with limited profit shifting was to mitigate tax-based distortions in the investment profile, the evaluation of the policy's impact on production efficiency is crucial to gauge its adequacy.

In the model, an inefficiency in the production process is captured by the aggregate additional capital costs of mobile firms. The analysis is focused on the impact in a mixed-strategy equilibrium and differences to the pure strategy equilibrium are clearly indicated throughout the discussion. Algebraically, the inefficiency in a mixed-strategy equilibrium is given by:

$$\mathrm{IE} = \Pr\left(\bar{t}\right)^2 m \int_0^1 \mathrm{ACC}^f df + m \int_{t'}^{t''} \left\{ \left[\int_0^{f^{thr}(t)} \mathrm{ACC}^f df \right] \Pr(t_{min} = t) \right\} dt$$

If both high productive countries charge their domestic monopoly rate, by the definition of mobility, all mobile firms will distort their investment profile in order to shift profits to the low tax-low productive country. Differently, if at least one productive country competes for mobile firms with a tax rate $t \in [t', t'']$, only a share $f^{thr}(t)$ of mobile firms incurs additional capital costs. The marginal inefficiency with respect to α is given by:

$$\begin{aligned}
\frac{\partial \mathrm{IE}}{\partial \alpha} = & \underbrace{m \int_{t'}^{t''} \frac{\partial f^{thr}(t)}{\partial \alpha} \mathrm{ACC}^{f^{thr}(t)} \Pr(t_{min} = t) dt}_{\text{extensive firm}} \\
& \underbrace{+ \Pr\left(\bar{t}\right)^2 m \int_0^1 \frac{\partial \mathrm{ACC}^f}{\partial \alpha} df + m \int_{t'}^{t''} \left[\int_0^{f^{thr}(t)} \frac{\partial \mathrm{ACC}^f}{\partial \alpha} df \right] \Pr\left(t = t_{min}\right) dt}_{\text{intensive firm}} \\
& \underbrace{+ 2 \frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha} \Pr\left(\bar{t}\right) m \left[\int_0^1 \mathrm{ACC}^f df - \int_0^{f^{thr}(t'')} \mathrm{ACC}^f df \right] - m \int_{t'}^{t''} \left\{ \frac{\partial \left[\int_0^{f^{thr}(t)} \mathrm{ACC}^f df \right]}{\partial t} \frac{\partial \Pr\left(t_{min} \leq t\right)}{\partial \alpha} \right\} dt}_{\text{strategic}}
\end{aligned}$$

The impact of the profit shifting limitation on inefficiency is carried by three separable effects. First, I formulate the Proposition regarding the total impact. Second, I precisely explain the channels of impact, provide intuition for each channel in detail and show how the channels add up to generate the total effect.

Proposition 4. inefficiency

(i) *If the limitation on profit shifting is tight, $\alpha \in [1, \alpha^{thr}]$, a lift in the limitation will increase the inefficiency.*

(ii) *If the limitation on profit shifting is weak, $\alpha > \alpha^{thr}$, a lift in the limitation will decrease the inefficiency.*

In a first step, to disentangle the total impact, initially keep tax rates and the mass of mobile firms investing in C, f^{thr}, fixed. The impact of a lift in the limitation on profit shifting on the inefficiency is now only due to the change in behavior of firms already investing in C, the inframarginal firms. In particular, inframarginal firms will allocate back capital to their more productive facilities in A, retaining the profit shifting ability. In consequence, this *intensive firm effect* decreases the additional capital costs of inframarginal firms, $\frac{\partial ACC}{\partial \alpha} < 0$, and so, decreases the global production inefficiency.

In a second step, keep tax rates fixed, but endogenize the mass of mobile firms investing in C, f^{thr}. The additional impact of a lift in the limitation on profit shifting on the inefficiency is now due to the change in behavior of firms, who were previously just about to invest in C, the marginal firms. In their location decision, marginal firms are more sensitive to the fixed tax incentive for investment in C, because the additional capital costs from investment decrease. This *extensive firm effect* can be interpreted as a standard price effect. Given that the return to investment (in terms of tax savings) in C increases, more firms allocate capital there, $\frac{\partial f^{thr}(t)}{\partial \alpha} > 0$. The extensive firm effect increases the global inefficiency, because marginal firms reduce their average capital productivity. Finally, the effect phases out in the level of α. If the profit shifting regulation is sufficiently weak, a vast majority of firms already has investment in the low productive-low tax country, so that the scope for the extensive firm effect decreases.

The effects presented so far are completely firm-based. A lift in the limit on profit shifting decreases the inefficiency, because it lowers the per-firm inefficiency. However, exactly because a lift in the limitation lowers the per-firm inefficiency, it incentivizes more firms to start generating an inefficiency. The relative magnitude of both effects structurally depends on the initial degree of limitation. The tighter the limit on profit shifting, the fewer firms already generate an inefficiency, the weaker the former effect. However, in my model, the intensive effect always dominates, so that taking into account firm-based effects only, the inefficiency monotonically decreases in the limitation on profit shifting. Finally, see that the qualitative results of the firm-based effects carry over for more general production technologies, which will be discussed in Section 6.

In a third step, endogenize the tax rates. The additional impact of a lift in the limitation on profit shifting on inefficiency is now due to the change in equilibrium tax rates. As a response to a lift in the limitation on profit shifting both types of countries, high and low productive, increase tax rates. By the more generous limit on profit shifting high productive countries lose the ability to leverage on their competitive advantage in capital productivity during the competition for mobile tax revenues against the low productive country. As a strategic response, high productive countries increase their tax rates to capture a larger share of the domestic monopoly tax revenues. This pull out of competition by high productive countries concentrates the supply side of the market for sheltering mobile firm profits. As a strategic response, the remaining supplier on the market, the low productive country, increases the tax rate due to the slacker competitive constraints. However, to balance out the elasticity, the low productive country increases the tax rate by less than unity, so that the positive difference between tax rates in high and low productive countries increases. That widening in the tax difference will incentivize more mobile firms to start allocating capital to low productive-low tax country and so, increases the inefficiency. That is the *strategic effect*.

Strategic effects increase the inefficiency by the selection of high productive countries out of competition, rather than by a pure concentration effect. Additional capital costs capture the competitive advantage of high productive countries on real economic grounds. Raising the limit on profit shifting lowers the additional capital costs. So, the competitive position of high and low productive countries in the competition for mobile firms is adversely affected by a change in the limitation. In particular, raising the limit worsens the position of high productive countries. Accordingly, the outside option of domestic monopoly revenues becomes relatively more attractive and high productive countries start adjusting their tax rates upward, closer to the domestic monopoly rate.

Strategic effects are only present for relatively tight profit shifting regulations. If the regulation is sufficiently weak, high productive countries already selected out of the competition for mobile firms. Then a further lift of the limitations on profit shifting does not affect the composition of the countries competing for mobile firms' tax revenues. If present the strategic effect dominates the firm-based effect and so, generates the non-monotonicity of the inefficiency in the degree of profit shifting regulation, as depicted in Figure 5.

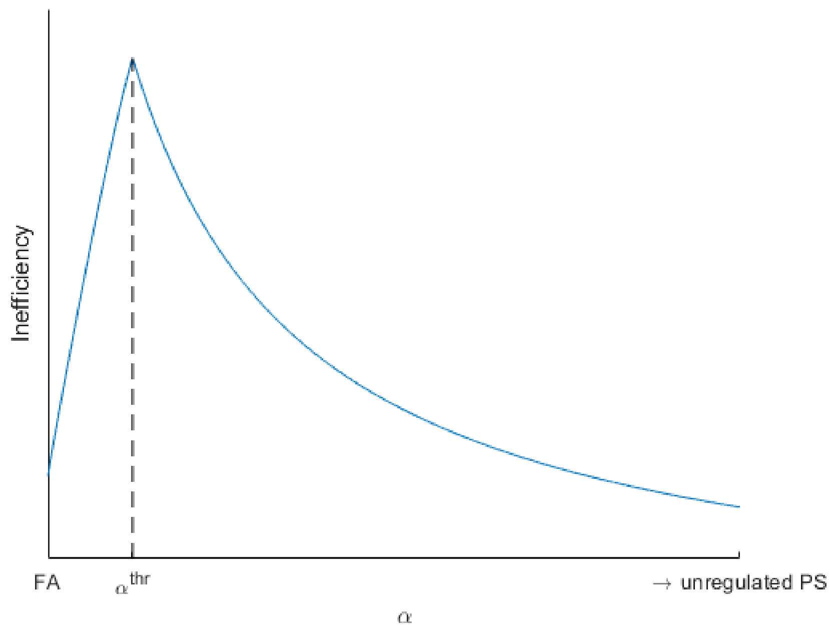


Figure 5: Inefficiency as a function of α

In sum, taking into account the intensive mechanic effects only, allowing for some limited degree of profit shifting seems to be a functioning tool to mitigate tax-motivated distortions in the investment profile. Since an alleviation of the apportionment factors reduces the tax "penalty" of allocating capital to productive but tax heavy jurisdiction, a firm seeking to minimize the tax bill increases the average productivity of its capital allocation. However, this effect is dampened by more firms investing in less productive countries, so that already the mechanic effect of limited profit shifting is ambiguous. At a parallel, the allowance for limited profit shifting attenuates the incentives for countries to participate in the competition for mobile firms. Generally this tendency of market concentration does not affect the inefficiency. However, given that the incentive to pull out from the competition is on countries with superior productivity, concentration decreases the average productivity of competing countries. Consequently, by the lack of productive alternatives, more firms are pushed into countries with inferior productivity. The marginal inefficiency of a profit shifting allowance increases by the strategic effects.

If a policymaker was interested in the containment of profit shifting and/or maximizing production efficiency, she should never implement a policy on the upward-sloping part of the inefficiency curve, a policy with $\alpha \in \left[1, \alpha^{thr}\right]$, rather than a tight Proportional Apportionment regime. The modified Nexus approach of the OECD features a limit on profit shifting of $\alpha = 1.3$. Further, the policy targets the competition for tax revenues from intangible assets by the provision of preferential tax rates for IP-revenues, so-called "patent box regimes". To date, the supply side of that market is characterized by a set of competitors, heterogeneous in productivity. Thus, the OECD's decision to allow for limited profit shifting can generate an increase in the inefficiency by the adverse effects on competition as described above.

4.2 Marginal Profits

According to the results of Section 4.1, the emergence of policies with limited profit shifting conflicts with their official rationale, production efficiency. So, this subsection provides a different rationale for the introduction of limited profit shifting policies, based on the distribution of marginal profits across mobile firms.

This section is focused on the marginal profits of mobile firms, if the limitation on profit shifting is tight, e.g. for $\alpha \in \left[1, \alpha^{thr}\right]$. The focus leads to the formulation of a competing rationale for the introduction of limited profit shifting policies for a situation, in which production efficiency is incapable of. In particular, consider the equilibrium post-tax profits of a mobile firm with a degree of mobility implied by f. Recall that $f \in [0,1]$ and that the higher f, the lower a firm's degree of mobility. In a mixed strategy-equilibrium, the individual profits of a mobile firm f read:

$$\text{MoP}^f = \Pr(\bar{t})^2 \Pi_L^f(t^m) + \int_{t'}^{t''} \overbrace{\left\{ \mathbb{I}_{f^{thr}(t) \geq f} \Pi_L^f\left(t_n(t)\right) + (1 - \mathbb{I}_{f^{thr}(t) \geq f}) \Pi_H(t) \right\}}^{\max\left\{\Pi_L^f(t_n(t));\Pi_H(t)\right\}} \Pr(t_{min} = t)$$

Here, $\Pi_{L[H]}^f(t)$ depicts the post-tax profits of firm f when paying a tax rate t in a low [high] productive country. Further, $\mathbb{I}_{f^{thr}(t) \geq f}$ indicates that, given $t_{min} = t$, firm f maximizes profits by investing in and shifting to the low productive country. By the definition of f^{thr}, selecting the maximum of the profit levels is equivalent to the expression using the indicator function. The marginal profits with respect to α of a mobile firm f are given by:

$$\frac{\partial \text{MoP}^f}{\partial \alpha} = \underbrace{- \left[\Pr(\bar{t})^2 \frac{\partial \text{ACC}^f}{\partial \alpha} + \int_{t'}^{t''} \left\{ \mathbb{I}_{f^{thr}(t) \geq f} \frac{\partial \text{ACC}^f}{\partial \alpha} \right\} \Pr(t_{min} = t) \right]}_{\text{intensive firm}} \tag{9}$$

$$\underbrace{+ 2 \frac{\partial \Pr(\bar{t})}{\partial \alpha} \Pr(\bar{t}) \left[\Pi_L^f(t^m) - \max\left\{ \Pi_L^f\left(t_n(t'')\right); \Pi_H(t'') \right\} \right] - \Pr(\bar{t})^2 \frac{\partial t^m}{\partial \alpha} \Pi - \int_{t'}^{t''} \frac{\partial \max\left\{ \Pi_L^f\left(t_n(t)\right); \Pi_H(t) \right\}}{\partial t} \frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha}}_{\text{strategic}}$$

The overall marginal profits can be split up in strategic and firm-based effects. The mechanisms behind the individual effects remain the same as in the analysis of 4.1. In general, firm-based and strategic effects point in opposite directions.

On the one hand, strategic effects lead to a rise in the economy's level of taxation. First, high productive countries select out of competition for mobile tax revenues by raising tax rates, closer to the domestic monopoly tax rate. Second, the low productive country remains on the market for mobile tax revenues and increases its tax rate, because it faces slacker competitive constraints. All together, the strategic effects decrease post-tax profits for each mobile firm in the economy.

On the other hand, the intensive firm effect leads to a decrease in additional capital costs for mobile firms investing in the low productive country. These inframarginal firms are allowed to fetch back parts of their investment to more productive facilities, retaining the same shifting allowance. So, the firm-based effect weakly increases post-tax profits for each mobile firm in the economy.

The distinction between strategic and firm-based effects outlines the general ambiguity of marginal profits from a lift in the profit shifting limitation. The rest of this section investigates on how the relative magnitude of both effects depends on a mobile firm's degree of mobility. Initially, consider the boundaries of the mobility support.

Lemma 3. the perfectly mobile firm
If the regulation allows for more profit shifting in the economy, profits of the perfectly mobile firm, $f = 0$, will unambiguously decrease.

The perfectly mobile firm has the same capital productivity in high and low productive countries. So, the ability to shift profits at a lower level of investment does not affect the capital costs, $\frac{\partial \text{ACC}^0}{\partial \alpha} = 0$, the firm-based effect does not affect marginal profits. At the same time, post-tax profits of the perfectly mobile firm decrease, because the low tax country increases its tax rate. For the perfectly mobile firm, the impact of the strategic effect is unambiguously stronger than the firm-based effect.

Lemma 4. the least mobile firm
If the regulation allows for more profit shifting in the economy, profits of the least mobile firm, $f = 1$, unambiguously decrease.

As long as high productive countries compete for mobile tax revenues, the least mobile firm will always decide to enjoy the better business environment and pay higher tax rates in exchange, $f^{thr}(t) < 1 \; \forall t \in [t', t'']$. Thus, the firm gains from the intensive firm effect realize on a mass market only. However, as a mass provider, the low productive country uses its market power to skim away the capital costs relief by increasing its tax rate accordingly. The strategic tax increase of the low productive country dominates the intensive firm effect, $\frac{\partial t^m}{\partial \alpha}\Pi \geq -\frac{\partial \text{ACC}^f}{\partial \alpha}$. Taking into account the strategic tax increase of the high productive country shows that for the least mobile firm, the impact of the strategic effect is unambiguously stronger than the firm-based effect.

For clarity, the results at the boundaries of the mobility support were shown to generally hold in a mixed strategy equilibrium. Instead, to discuss the general impact on firms, the focus is explicitly on policy changes away from a tight Proportional Apportionment, i.e. the focus is on a marginal increase in α at $\alpha = 1$. Like that, we analysis addresses the question why policymakers have introduced limited profit shifting policies instead of standard Proportional Apportionment.

Proposition 5. a mobile firm
Starting at a Proportional Apportionment regime, $\alpha = 1$, marginal profits from a lift in the profit shifting limitation will be highest for an intermediate firm $f^* \in (0, 1)$.

As a strategic response to a lift in the limitation on profit shifting, high productive countries increase tax rates to extract a larger share of the domestic monopoly tax revenues. In turn, the low productive country increases its tax rate due to the slacker competitive constraints on the market for tax revenues from mobile firms. However, in order to balance out the tax base elasticity, the low productive country increases the tax rate by less than the high productive countries. Accordingly, the negative impact of a more generous profit shifting regulation on marginal profits through an increase in the level of taxation (the strategic effect) increases in the (expected) amount of production in the low productive country. The (expected) amount of production in the low productive country increases in mobility. Thus, the negative strategic effect monotonically increases in a firm's degree of mobility. Putting it differently, in response to a lift in the limitation on profit shifting, the (expected) average tax rate of firms with a low degree of mobility increases the most.

Thus, the non-monotonicity in marginal profits with respect to a firm's degree of mobility stems from the firm-based effect. In fact, the intensive firm effect is governed by two conflicting forces.

On the one hand, the positive impact through the intensive firm effect increases in the (expected) amount of investment in the low productive country. After a lift in the limitation on profit shifting, each capital unit invested in the low productive country generates a higher post-tax profit for the mobile firm. The profits realize either because a mobile firm allocates back the capital to decrease its capital costs, or because a firm keeps the capital in the low productive country at a higher shifting per investment rate. Firms with a high degree of mobility have the highest (expected) amount of capital investment in the low productive country. Thus, *by quantity*, the positive intensive firm effect increases in mobility.

On the other hand, the positive impact through the intensive firm effect increases in the additional capital costs per unit. Accordingly, the value of allocating a single unit of capital back to the high productive country decreases in a firm's degree of mobility. Thus, *by unit*, the positive intensive firm effect decreases in mobility.

Firms with an intermediate degree of mobility balance out by quantity- and per unit-effect and so, gain the most through the intensive firm effect. For an individual firm, the intensive firm effect is first order and so the shape of the total effect follows the intensive firm effect, as depicted in Figure 6. In particular, these firms have a significant share of production in the low productive country and generate a substantial inefficiency per unit of investment.

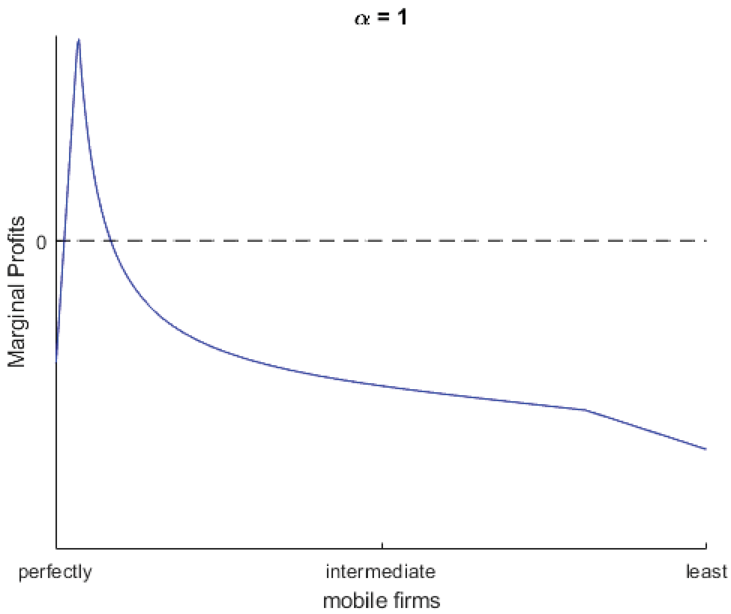


Figure 6: Marginal Profits as a function of mobility

Marginal profits from a lift in the limitation on profit shifting are not symmetrically distributed across firms, but are concentrated on a subset of firms. Based on this result, the rest of this section formulates a novel rationale for the emergence of policies with limited profit shifting, different from the enhancement of production efficiency.

In theories of lobbying, oftentimes results from the literature on collective action have been applied to predict the probability of success for a certain lobbying agenda. An important insight from that literature, mainly influenced by Olson (2009), is that a collective action is most likely to be implemented, if each individual's contribution is pivotal for the overall success of the action. In such situations, the free-rider problem is mitigated and the action that maximizes welfare for the collective is likely to be chosen.

Applying this logic to the regulation of profit shifting, it seems that limited profit shifting policies provide a breeding ground for lobbying activities. Gains from limited profit shifting are concentrated on a subgroup of agents, however, losses from such policy more evenly spread out over a broader mass of firms. Thus, one could expect the winning firms to more easily coordinate on a collective voice in the legislative process of limited profit shifting policies, like the OECD modified Nexus. In fact, the initial Nexus approach of the OECD followed a standard Proportional Apportionment, $\alpha = 1$. Only in a second draft, the OECD modified the approach by the allowance for some limited profit shifting, choosing an uplift parameter of $\alpha = 1.3$. As pointed out earlier, the official rationale was the mitigation of tax-based distortions in the international allocation of production. However, during the legislative process, firms have been vocal, so that the introduction of a policy with limited profit shifting 'reflect(s) these concerns raised by businesses.' OECD (2015). Finally, NGOs like the BEPS Monitoring Group have already earmarked the introduction of limited profit shifting as a succesful lobbying activity: 'We believe that the agreed 30% uplift approach is merely another give-away that is simply not appropriate.' (BEPS Monitoring Group, 2015).

In sum, this subsection has provided a novel rationale for the introduction of limited profit shifting policies. The theoretical idea behind this rational, gains from limited profit shifting are concentrated on a limited subgroup of agents, rely on clean economic arguments. The interpretation of these results, concentrated gains provide a breeding ground for lobbying activities, has to be read cautiously.

5 The effect on other economic outcomes

This section complements the previous section with results on other economic outcomes of interest. In particular, this section considers the impact of the limitation on profit shifting on national income of countries and aggregate profits of firms in the economy.

5.1 High productive countries

In this subsection, I consider the impact of the limitation on profit shifting on the national income of the high productive countries, which is defined as the sum of tax revenues and post-tax profits of the domestic firms. The following Proposition summarizes the main results of this section:

Proposition 6. high productive countries - national income
If the limit on profit shifting is tight, $\alpha \in \left[1, \alpha^{thr}\right]$, national income will decrease in the limitation on profit shifting. If the limit on profit shifting is weak, $\alpha > \alpha^{thr}$, national income will not change.

In absence of mobile tax revenues, the national income of a high productive country is given by the profits of its domestic firms, $d(\Pi - \frac{r}{\theta}F) \equiv \mathrm{I}^{\mathrm{aut}}$. The level of the tax rate determines the split of $\mathrm{I}^{\mathrm{aut}}$ in private profits and public revenues, but is irrelevant for the level of national income. In a pure strategy equilibrium, a high productive country taxes its domestic firms only and thus, national income is $\mathrm{I}^{\mathrm{aut}}$, independent of the limitation on profit shifting. On the other hand, in a mixed-strategy equilibrium (in expectations) a high productive country attracts mobile tax revenues and thus, national income is higher than $\mathrm{I}^{\mathrm{aut}}$. In particular, the expected national income is given by:

$$\mathrm{I}^{\mathrm{mix}} = \underbrace{\left[\Pi - \frac{r}{\bar{\theta}} F\right] d}_{\mathrm{I}^{\mathrm{aut}}} + \int_{t'}^{t''} mt \left(1 - f^{thr}(t)\right) \Pi \Pr(t_{min} = t) \tag{10}$$

The second term of (10) represents the expected revenues from the mobile tax base. It sums across the tax rate support the mobile revenues from competing for the mobile firms, $mt(1-f^{thr}(t))\Pi$, multiplied by the probability of underbidding the other high productive country at tax rate t, $\Pr(t_{min} = t)$. During the construction of a mixed-strategy equilibrium, it was shown that for every t, expected tax revenues as given by (8) have to equal the domestic monopoly tax revenues. Using this identity, (10) reformulates:

$$\mathrm{I}^{\mathrm{mix}} = \mathrm{I}^{\mathrm{aut}} + \int_{t'}^{t''} \left(\bar{t} - t\right) d\Pi g(t) \tag{11}$$

The difference between national income levels, $\mathrm{I}^{\mathrm{mix}} - \mathrm{I}^{\mathrm{aut}}$, is given by the expected tax relief for domestic firms. By taxing part of the mobile firms, a high productive country is able to take away tax burden from the domestic producers and thus, increases national income. The impact of the limitation on profit shifting on national income is algebraically described by:

$$\frac{\partial \mathrm{I}^{\mathrm{mix}}}{\partial \alpha} = d\Pi \left[-\left(\bar{t} - t''\right) \frac{\partial Pr\left(\bar{t}\right)}{\partial \alpha} + \int_{t'}^{t''} \frac{\partial G(t)}{\partial \alpha} \right] < 0$$

An increase in α triggers a First-Order Stochastic dominant shift in the mixing probabilities. If more profit shifting is allowed in the economy, high productive countries gradually pull out of the competition for the mobile firms. By attenuating the additional capital costs of investment in a low productive country, a higher level of α diminishes the ability of a high productive country to generate tax revenues from the competition for mobile firms. Accordingly, a higher limit on profit shifting spurs the selection of high productive countries out of the competition for mobile tax revenues. This way, the tax burden is gradually shifted to the domestic firms until at some level of α, in the model α^{thr}, competition for mobile tax revenues stops and domestic firms have to carry the full tax burden.

5.2 Low productive country

Given that the domestic tax base of low productive countries is of negligible size, national income is equivalent to mobile tax revenues. This subsection analyzes the impact of the limitation on profit shifting on mixed-strategy equilibrium tax revenues of the low productive country. However, differences to the pure strategy equilibrium are clearly indicated throughout the discussion. Algebraically, the mixed-strategy equilibrium tax revenues are given by:

$$\mathrm{TR} = \overbrace{mt^m \Pi}^{TR^m} \Pr\left(\bar{t}\right)^2 + \int_{t'}^{t''} \overbrace{mt^n(t) f^{thr}(t) \Pi}^{TR^n(t)} \Pr(t_{min} = t) \tag{12}$$

Tax revenues of the low productive country structurally depend on the number of competing countries in the economy. If none of the high productive countries competes for tax revenues from mobile firms, the low productive country acts as a mass provider for sheltering the profits of mobile firms. With at least one high productive country competing, the low productive country chooses its tax rate to serve only a fraction of the mobile firms. In a pure strategy equilibrium, the low productive country always becomes a mass provider, whereas in a mixed-strategy equilibrium the probability of being a mass provider is given by $Pr(\bar{t})^2 < 1$. The marginal tax revenues from a lift in the limitation on profit shifting can be segregated in three interpretable terms:

$$\frac{\partial \mathrm{TR}}{\partial \alpha} = \underbrace{\Pr\left(\bar{t}\right)^2 \frac{\partial t^m}{\partial \alpha} m\Pi}_{\text{strategic low}} + \underbrace{\int_{t'}^{t''} m t^n(t) \frac{\partial f^{thr}(t)}{\partial \alpha} \Pi \Pr\left(t_{min} = t\right)}_{\text{extensive firm}} \tag{13}$$
$$+ \underbrace{2\frac{\partial \Pr(\bar{t})}{\partial \alpha} \Pr(\bar{t}) \left(TR^m - TR^n(t'')\right) - \int_{t'}^{t''} \frac{\partial TR^n(t)}{\partial t} \frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha}}_{\text{strategic high}}$$

With at least one high productive country competing for tax revenues from mobile firms, there exists a marginal firm indifferent to investing capital in and shifting profits to the low productive country. A lift in the limitation on profit shifting attenuates the additional capital costs of investment and thus, the marginal firm starts claiming profits in the low productive country. The impact of this relocation on tax revenues is captured by the *extensive firm effect*. The effect describes the mechanic marginal revenues for a low productive country.

As a mass provider, the low productive country optimally chooses a tax rate to capture full profits of the least mobile firm. Accordingly, the strategic response of the low productive country to a lift in the limitation on profit shifting is to skim away the capital costs relief of the least mobile firm by increasing the tax rate accordingly, labeled as *strategic low effect*.

Finally, the strategic response of high productive countries is to focus more on taxation of the domestic firms by shifting probability mass to higher realizations of t. On the one hand, the distributional shift increases the probability of the low productive country being the sole competitor on the market for tax revenues from mobile firms, $\frac{\partial \Pr(\bar{t})}{\partial \alpha} > 0$. On the other hand, the distributional shift increases the (expected) tax rates of high productive countries, $\frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha} < 0$. The budgetary impact through the strategic response of high productive countries is labeled as *strategic high effect*.

Proposition 7. low productive country - national income

(i) *The tax revenues of the low productive country mechanically increase in the limitation on profit shifting.*

(ii) *Taking into account the strategic marginal revenues amplifies the positive mechanic effect.*

Regarding (i), the extensive firm effect increases the tax revenues of the low productive country as a positive tax base effect for niche providers, $\frac{\partial f^{thr}(t)}{\partial \alpha} > 0$. In reference to (ii), the strategic marginal revenues are twofold. On the one hand, the strategic low effect increases revenues as a positive tax rate effect on an inelastic tax base for mass providers, $\frac{\partial t^m}{\partial \alpha} > 0$. On the other hand, the strategic high effect increases revenues by shifting outward the expected demand for profits claimed in the low productive country, $\frac{\partial \mathrm{TR}^n(t)}{\partial t} > 0$. Strategic revenues are unambiguously positive, amplifying the positive mechanic impact. Finally, in a pure-strategy equilibrium high productive countries already selected out of competition, so that the strategic low effect is the only channel of impact. Thus, in a situation without competing high productive countries, tax revenues of the low productive country increase by the pure strategic low effect.

In summary, the low productive country benefits from a more lenient profit shifting regulation in the economy. Here, strategic effects reinforce the mechanic impact on tax revenues. Thus, embedding the profit shifting regulations in a more general tax competition framework does not change the evaluation for low productive countries qualitatively, but works out additional channels with positive impact. In particular, diminishing incentives for high productive countries to participate in the competition for the mobile tax base as well as slacker competitive constraints imposed by non-competing countries increase tax revenues in addition.

5.3 Mobile Firms

This subsection analyses the impact of the profit shifting regulation on the aggregate profits of the mobile firms. Section 4.2 has analyzed the marginal profits on a micro-level. This subsection sums up the individual effects to determine the macro-level impact on profits of the mobile firms. The analysis is focused on marginal profits in a mixed-strategy equilibrium. Differences to the pure-strategy equilibrium are clearly indicated throughout the discussion.

$$\mathrm{MoP} = \Pr\left(\bar{t}\right)^2 \overbrace{m \int_0^1 \Pi_L^f\left(t^m\right) df}^{\mathrm{MoP}^m} + m \int_{t'}^{t''} \left\{ \overbrace{\left[\int_0^{f^{thr}(t)} \Pi_L^f\left(t^n(t)\right) df + \left(1 - f^{thr}(t)\right) \Pi_H(t) \right]}^{\mathrm{MoP}^n(t)} \Pr(t_{min} = t) \right\} dt$$

As an example, $\Pi_L^f(t^m)$ describes the post-tax profits of a firm with productivity indexed by f, claiming profits in the low productive country at a tax rate of t^m. If the firms do not allocate capital to the low productive country, profits are symmetric across firms, $\Pi_H^f = \Pi_H$. However, for firms investing in the low productive country, profits are decreasing in index f. The larger f, the less productive a unit of capital is in the low productive country, the higher the additional capital costs to meet the requirement for a shift of profits. For the ease of exposition, I denote by $\mathrm{MoP}^n(t)$ (MoP^m) the aggregate profits of mobile firms in an environment with multiple competing countries (a sole competitor).

$$\frac{\partial \mathrm{MoP}}{\partial \alpha} = \underbrace{-\Pr(\bar{t})^2 m \frac{\partial t^m}{\partial \alpha} \Pi}_{\text{strategic low}} - \underbrace{\left[\Pr(\bar{t})^2 m \int_0^1 \frac{\partial \mathrm{ACC}^f}{\partial \alpha} df + \int_{t'}^{t''} \left\{ \left[\int_0^{f^{thr}(t)} \frac{\partial \mathrm{ACC}^f}{\partial \alpha} df \right] \Pr(t_{min} = t) \right\} dt \right]}_{\text{intensive firm}}$$
$$\underbrace{+ 2 \frac{\partial \Pr(\bar{t})}{\partial \alpha} \Pr(\bar{t}) \left[\mathrm{MoP}^m - \mathrm{MoP}^n(t'')\right] - m \int_{t'}^{t''} \left\{ \frac{\partial \mathrm{MoP}^n(t)}{\partial t} \frac{\partial \Pr(t_{min} \le t)}{\partial \alpha} \right\} dt}_{\text{strategic high}}$$

The mechanisms underlying the individual effects were extensively explained in subsection 4.2 already. The main results of this subsection are provided in the following Proposition:

Proposition 8. mobile firms - aggregate profits

(i) *If the limitation on profit shifting is lifted, the aggregate profits of mobile firms mechanically increase.*

(ii) *If the limitation on profit shifting is lifted, the aggregate profits of mobile firms decrease due to strategic effects. The strategic effects outweigh the positive mechanic effect; the profits of mobile firms overall decrease, if the limitation on profit shifting is lifted.*

Regarding (i), the marginal mechanic profits are given by the intensive firm effect. The allowance to shift more profits grants firms more flexibility in the capital allocation. As a response, inframarginal firms reallocate investment to their more productive facilities, still shifting the full amount of profits to the less productive country. The associated increase in average capital productivity reduces the global demand for capital and so additional capital costs for all mobile firms decrease, $\frac{\partial \text{ACC}^f}{\partial \alpha} \leq 0 \ \forall f \in [0, 1]$. Thus, aggregate profits of the mobile firms increase by the mechanic effect.

Regarding (ii), the marginal strategic profits are twofold. First, if the low productive country serves the mobile firms as a mass provider, a lift in the limit on profit shifting relaxes the competitive constraint from high productive countries outside the market. As an optimal response, the low productive country raises its tax rate, $\frac{\partial t^m}{\partial \alpha} > 0$, increasing the tax costs of the mobile firms, captured by the strategic low effect. Second, as a response to a lift in the limitation on profit shifting, high productive countries shift probability mass to higher tax rates. On the one hand, the shift increases the probability of a mass provider, $\frac{\partial \text{Pr}(\bar{t})}{\partial \alpha} > 0$. Due to the lack of direct competition, on a mass market the low productive country is able to generate a higher per-firm profit margin. Consequently the centralization tendencies decrease profits of the mobile firms, $\text{MoP}^m < \text{MoP}^n$. On the other hand, the effect increases the expected tax rates of high productive countries, $\frac{\partial \text{Pr}(t_{min} \leq t)}{\partial \alpha} < 0$. As tax rates are strategic complements, the aggregate profits of the mobile firms decrease in the tax rate of high productive countries, $\frac{\partial \text{MoP}^n(t)}{\partial t} < 0$. In a pure-strategy equilibrium, the willingness of high productive countries to compete for tax revenues of mobile firms has vanished already, such that marginal profits are determined by strategic low and intensive firm effect only, keeping the same intuition. Independent of the type of equilibrium, it can be shown that the strategic effect(s) always dominate and thus, a lift in the limitation on profit shifting decreases the aggregate profits of mobile firms, against the mechanic intuition.

In summary, if evaluating the impact of a lift in the limitation on profit shifting on the aggregate profits of mobile firms, strategic effects point in the opposite direction of the mechanic effect. Thus, evaluating the profit shifting regulation in a more general tax competition framework enriches the analysis qualitatively. While the mechanic effect increases the firms' profits by reducing the additional capital costs of a distorted capital allocation, diminishing incentives for high productive countries to participate in the competition for the mobile tax revenues as well as slacker competitive constraints imposed by non-competing countries increase the tax burden of mobile firms. While the overall sign of the impact is due to the parametrization of the model, the trade-off worked out in this subsection is of a more general nature.

6 Extensions

6.1 Competitive Sequence

The model relies on two types of specific features. On the one hand, I impose Stackelberg-leadership of high productive countries. On the other hand, competition between high productive countries is modelled as a duopoly, whereas the low productive country is the sole competitor on its stage. In this subsection I will discuss sensitivity of my qualitative results and provide intuition about possible extensions to a more general framework.

First, consider the sequential set-up of the model. I argue that the model's qualitative results are not sensitive to the particular competitive sequence. To see this, flip the competitive sequence, i.e. the low productive country chooses its tax rate on the first stage, high productive countries choose their tax rates as best responses. In such an environment, high productive countries compete against each other, given a (non-best response) tax rate of the low productive country. Accordingly, the difference between tax rates in high and low productive countries increases in a high productive country's tax rate. The model still generates a negative tax rate elasticity of the mobile tax base for high productive countries. In consequence, the properties of the equilibrium remain the same as in the exposition of this paper. Regarding a lift in the limitation on profit shifting, the main qualitative mechanisms continue to hold. If the limitation on profit shifting is lifted, given a low productive country's tax rate, high productive countries shift probability mass to higher tax rates. In consequence, the negative elasticity of demand for the low productive country will increase. As a strategic response, the low productive country increases its tax rate upward as well, however, the (expected) difference between tax rates increases to balance the elasticity at equilibrium. Thus, despite thresholds and cut-offs might change, the general equilibrium properties and the effects on economic outcomes will continue to hold under the alternated competitive sequence. I expect the same to hold in a simultaneous move game between low and high productive country, however, more work needs to be done to investigate on this.

Second, consider the intra-stage competition of the model. Throughout the model, a low productive country is the sole competitor on its stage. This modelling choice is made to provide a lower bound for the effect of a lift in the limitation on profit shifting on production efficiency. The necessary condition for an increase in the inefficiency due to strategic effects is that the (expected) difference between tax rates in high and low productive countries increases in response to a lift in the limit on profit shifting. In this regard, modelling a sole low productive country mitigates the strategic effect, because the low productive country uses its market power to increase the tax rate, and so attenuates the increase in the (expected) difference in tax rates. Instead, consider a situation with multiple homogeneous low productive countries on the second stage. By standard Bertrand-competition, equilibrium tax rates of low productive countries would equal zero in such a model, independent of the limitation on profit shifting. Accordingly, the impact of the strategic effect on inefficiency would be amplified. At the same time the impact of strategic effects on marginal profits of the mobile firms would change. In fact, given that low productive countries do not adjust tax rates, marginal profits for all firms would increase. For the perfectly mobile firm, marginal profits would be exactly zero in this case. However, the non-monotonicity of marginal profits and the result that only a subset of firms gains from the introduction of limited profit shifting would prevail.

6.2 Production technology

The model assumes a linear production technology. The modelling choice was motivated in Section 2.5, now I discuss sensitivity of the results to this simplifying assumption.

The firm-based effects in this model are the extensive margin- and intensive margin effect. In response to a lift in the limitation on profit shifting, marginal firms install more capital units in the low productive country and inframarginal firms allocate capital units back to the more productive facilities. Those effects would continue to hold under more general production techonologies.

Consider a situation, in which firms produce with respect to a production function $\mathrm{F}_j\left(\mathrm{K}_j\right)$ in each country. Assume that the production function meets the lnada conditions. No firm would ever centralize production in a certain country, but split the production according to its relative productivity levels and tax incentives. In difference to the exposition in this paper, the extensive firm effect would now capture the behavioral effect of a firm to allocate more capital units to the low productive country, instead of new firms starting their investment. Accordingly, in such a framework, for tight regulations the extensive firm effect would be relatively stronger as compared to the intensive firm effect. However, for weak regulations the intensive firm effect would take over again. So, the type of firm-based effects in the model carry over to more general production technologies. The strategic effects are qualitatively implied by the firm-based effects, so that the qualitative results are invariant to the specification of the production technology.

6.3 The modified Nexus approach

The modified Nexus approach (MNA) was formulated by the OECD during the recent initiative against Base Erosion and Profit Shifting. In particular, the MNA is part of Action 5 'Countering Harmful Tax Practices more effectively, taking into account Transparency and Substance'.

The regulation applies to the taxation of business profits from the provision of Intangible Property (IP) at preferential (reduced) tax rates. Over the last decade, several countries initialized preferential tax regimes for IP profits, so-called 'patent boxes'. The official rationale for the introduction of patent boxes is to stimulate domestic Research and Development. However, research over the last years has shown that patent boxes are mainly a tool for multinational enterprises to avoid taxation on IP profits. Accordingly, the OECD has identified patent boxes without an established link (nexus) between economic activity and preferential taxation as a harmful tax practice. In this regard, the introduction of the MNA is a necessary step for patent box countries to offer preferential tax rates on IP revenues without being blacklisted and so, subjected to counter-measures of other countries, like the imposition of withholding taxes.

The functional form of the MNA is directly adopted in this project. As previously mentioned, the OECD's uplift parameter is $\alpha = 1.3$. However, the practical implementation of the MNA differs from the exposition in my project in some ways. I will briefly introduce the two main differences.

First, in reality the limitation on profit shifting is a choice variable of the patent box country. In fact, within the OECD, the countries have agreed on an upper bound for the uplift parameter, $\overline{\alpha} = 1.3$. Each patent box country is allowed to choose its country-specific uplift parameter from the set $[1, 1.3]$. In a previous version of my model, countries were allowed to choose the parameter. The model predicts that low productive countries always choose $\alpha = \overline{\alpha}$ in equilibrium. This result is in line with recent observations on patent box countries like Belgium or Luxembourg.

Second, in reality the MNA apportions profits with respect to the allocation of 'qualified expenditures', not investment units. In my model, qualified expenditures would be the capital costs incurred within a certain country. Using expenditures instead of production would complicate the analysis significantly, however, not change any of my qualitative results. The trade-off in the capital allocation problem for firms would still exist; an investment in a low productive-low tax country triggers additional capital costs.

7 Conclusion

This paper has analyzed a rich model of tax competition with limited profit shifting. The first contribution of the project is to evaluate the impact of limited profit shifting on production efficiency. The results show that a tighter limit on profit shifting does not necessarily decrease production efficiency. Instead, the effect depends on the initial degree of limitation. In particular, tight regulations like Proportional Apportionment contain profit shifting at a locally maximized production efficiency. The intuition behind comes from the presence of strategic effects, which have so far been overlooked in the discussion on limit profit shifting policies.

The second contribution of this paper is to show that gains from allowing for some profit shifting are concentrated on a subset of firms with intermediate mobility. Highly mobile firms shift profits at negligible costs, so do not gain from the enhanced flexibility in profit shifting. Firms with a low degree of mobility never decide to set up multinational group structures to avoid taxation regardless and therefore, neither gain from a change in the regulation.

The results have policy relevance, because they challenge the widespread view that the fight against profit shifting necessarily faces a containment-efficiency trade-off. Further, the introduction of limit profit shifting policies has raised concerns about the role of business representatives in the legislative process. These concerns are featured by the result that marginal profits from the policy are concentrated on a subset of agents and so, make a collective lobbying effort more likely.

As the introduction of limited profit shifting policies, like the modified Nexus approach, is recent, it is too early to seek for complete empirical evidence. In the meantime, a potential extension of this project is to look for evidence for parts of the mechanisms identified in this paper. First, for profit shifting regulations of different tightness, do we observe the extensive- and intensive margin firm-based effects? Second, for profit shifting regulations of different tightness, do we observe different participation rates of high productive jurisdictions in competition?

A Proofs of Lemmas and Propositions

Proof of Lemma 1

I will show that conditional on investing in C being the optimal capital allocation, a mobile firm f will always invest such that $K_C^f = \frac{F}{\alpha\theta_C^f} = K_f^*(C)$. From (2) it follows that

$$\Pi_f(K_f^*(C)) = (1-t_C)\Pi - r\left[\frac{1}{\theta_C^f} + \frac{1}{\overline{\theta}}(\alpha - 1)\right]\frac{F}{\alpha}$$

Where the assumption is that f invests the additional capital in A to just meet the production constraint. Suppose $K_f^*(C)$ does not maximize post-tax profits and instead consider $K_f'(C) > K_f^*(C)$. The profits are given by:

$$\Pi_f(K_f'(C)) = (1-t_C)\Pi - r\left(K_f'(C) + \frac{F - \theta_C^f K_f'(C)}{\overline{\theta}}\right)$$

$K_f'(C)$ yields higher profits than $K_f^*(C)$, iff $\Pi_f(K_f'(C)) \geq \Pi_f(K_f^*(C))$, which boils down to $K_f'(C) \leq \frac{F}{\theta_C^f\alpha}$ and thus, contradicts the definition of $K_f'(C)$. Second, consider a $K_f'(C) < K_f^*(C)$. The profits are given by:

$$\Pi_f(K_f'(C)) = \left(1 - \left(1 - \alpha\frac{\theta_C^f K_f'(C)}{F}\right)t_A - \alpha\frac{\theta_C^f K_f'(C)}{F}t_C\right)\Pi - r\left(K_f'(C) + \frac{F - \theta_C^i K_f'(C)}{\overline{\theta}}\right)$$

A firm decides to invest $K_f'(C)$, only if $\Pi_f(K_f'(C)) > (1-t_A)\Pi - \frac{r}{\overline{\theta}}F$; profits from investing in C exceed profits of a centralized investment in A. The condition boils down to:

$$0 > K_f'(C)\left[r\left(1 - \frac{\theta_C^f}{\overline{\theta}}\right) + \alpha\frac{\theta_C^f}{F}(t_C - t_A)\Pi\right] \tag{A.1}$$

Further, f decides to invest $K_f'(C)$, only if $\Pi_f(K_f'(C)) \geq \Pi_f(K_f^*(C))$, which boils down to:

$$\left[r\left(1 - \frac{\theta_C^f}{\overline{\theta}}\right) + \alpha\frac{\theta_C^f}{F}(t_C - t_A)\Pi\right]K_f'(C) \leq \frac{F}{\theta_C^f\alpha}\left[r\left(1 - \frac{\theta_C^f}{\overline{\theta}}\right) + \alpha\frac{\theta_C^f}{F}(t_C - t_A)\Pi\right] \tag{A.2}$$

Given that $K_f'(C) > 0$, the in-bracket term of (A.1) has to be strictly smaller than zero. Applying this result to (A.2) generates $K_f'(C) \geq \frac{F}{\theta_C^f\alpha}$, which contradicts the initial assumption. Thus, if the firm decides to invest in C, it has to be with $K_f^*(C)$. Finally, the proof for a firm investing in B or A and B aside is symmetric and omitted for brevity.

Proof of Lemma 2

(i): Lemma 1 has shown that a firm pays taxes in A, only if it does not invest in C. So, profits in a situation with tax burden in A and B are given by:

$$\Pi(A\&B) = \left(1 - \left(1 - \min\left\{1; \alpha\frac{\overline{\theta}K_f(A)}{F}\right\}\right)t_B - \min\left\{1; \alpha\frac{\overline{\theta}K_f(A)}{F}\right\}t_A\right)\Pi - \frac{r}{\overline{\theta}}F \tag{A.3}$$

The respective level of capital investment in B is implied by the production constraint $K_f(A) + K_f(B) = \frac{F}{\overline{\theta}}$. See that capital costs are independent of the capital allocation, because mobile firms have a symmetric capital productivity in A and B. Finally, it is easy to show that for $t_B < t_A$, $K_f(A) = 0$ maximizes (A.3).

(ii) Using the results from Lemma 1 and (i) a mobile firm f invests in C, iff:

$$(1 - t_C)\,\Pi - r\left[\frac{1}{\theta_C^f} + \frac{1}{\overline{\theta}}(\alpha - 1)\right]\frac{F}{\alpha} > (1 - \min\{t_A; t_B\})\,\Pi - \frac{r}{\overline{\theta}}F \tag{A.4}$$

Thus, a necessary and sufficient condition for a positive mass of firms in C is that (A.4) holds for $f = 1$, i.e. that the perfectly mobile firm invests and pays taxes in C. The inequality boils down to $t_C < \min\{t_A; t_B\}$. Thus, a positive mass of mobile firms invests and pays taxes in C, if only if $t_C < \min\{t_A; t_B\}$.

Proof of Proposition 1

The first part of Proposition 1 is a direct adoption of the joint statements in Lemmas 1 and 2. For the second part, see that under the conditions of Proposition 1, firm f invests in C, iff $\Pi\left(K_f^*(C)\right) > (1 - t_A)\Pi - \frac{r}{\overline{\theta}}F$. The condition boils down to:

$$f < \alpha\frac{(t_A - t_C)}{\Delta}\frac{\Pi}{rF}$$

Given $t_C < t_A$ it unambiguously holds that $f > 1$, however, $f \leq 1$ is not guaranteed and thus, $f^{thr} = \min\left\{1; \alpha\frac{(t_A - t_C)}{\Delta}\frac{\Pi}{rF}\right\}$. As this shows the derivation of the threshold, the investment levels on the intensive margin were derived in the proofs of Lemmas 1 and 2.

Proof of Proposition 2

Using the distributional assumption as well as the results from Proposition 1, tax revenues are given by:

$$TR^C = \begin{cases} t_C m^{\frac{\frac{\alpha}{r}\frac{\Pi}{F}(t_{min} - t_C)}{\Delta}}\Pi & \text{if} \quad t_C \geq t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha} \\ t_C m\Pi & \text{else} \end{cases}$$

It can be shown that tax revenues are strictly concave on $t_C \in [t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha}, \overline{t}]$, linearly increasing on $t_C \in [0, t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha})$ and continuous over the whole interval $t_C \in [0, \overline{t}]$. Taking the First Order Condition piecewise, the optima on the subintervals are given by $t_C^{*up} = 0.5t_{min}$ for the upper piece and the corner solution $t_C^{*low} = t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha}$ for the lower piece. Given that revenues are increasing on the lower piece, concave on the upper piece and both pieces append by continuity, the global optimum on $\left[0, \overline{t}\right]$ is given by the global optimum on the upper piece, t_C^{*up}, if and only if $t_C^{*up} \geq t_C^{*low}$. Algebraically the condition derives as:

$$\begin{aligned} 0.5t_{min} &\geq t_{min} - \frac{rF}{\Pi}\frac{\Delta}{\alpha} \\ \leftrightarrow \quad t_{min} &\leq 2\frac{rF}{\Pi}\frac{\Delta}{\alpha} \end{aligned}$$

The derivation of the cut-off together with the particular policies completes the proof of Proposition 2.

Proof of Proposition 3

I proceed in three steps: (i) proof of existence of a unique equilibrium with $\alpha \geq \alpha^{thr}$ (ii) proof of non-existence of a pure-strategy equilibrium with $\alpha < \alpha^{thr}$ (iii) construction of a mixed-strategy equilibrium with $\alpha < \alpha^{thr}$

(i)

Consider a situation in which $t_B = \bar{t}$ and suppose that $\alpha \geq \alpha^{thr}$. By the choice of the parameter space, $t_A = \bar{t}$ yields certain tax revenues of $d\bar{t}\Pi$. In deviation, $t_A < \bar{t}$ yields the tax revenues from the competition with C as a Stackelberg leader. Proposition 2 has shown that the best response of C will be either $t_C^{*up} = 0.5t_A$ or $t_C^{*low} = t_A - \frac{rF}{\Pi}\frac{\Delta}{\alpha}$, depending on the choice of t_A. In particular, t_C^{*low} implies that the full set of mobile firms is moving to C. Thus, choosing a $t_A < \bar{t}$ such that the optimal response of C is given by t_C^{*low} is strictly dominated by $t_A = \bar{t}$. So, consider the tax revenues of A under the conjecture that the Stackelberg-follower responds with t_C^{*up}:

$$TR_A = t_A \left[d + m \left(1 - \alpha \frac{\Pi}{rF} \frac{0.5t_A}{\Delta} \right) \right] \Pi$$

For t_A such that $t_C = t_C^{*up}$ tax revenues are differentiable and strictly concave in t_A. So, using standard optimization techniques, the optimal tax rate is given by: $t_A^* = (1 + \frac{d}{m})\frac{rF}{\Pi}\frac{\Delta}{\alpha}$. It is easily verifiable that t_C^{*high} is indeed the best response to t_A^*. Further, the respective tax revenues are given by $TR_A^* = 0.5\frac{(d+m)^2}{m}\frac{rF}{\Pi}\frac{\Delta}{\alpha}\Pi$. Imposing the tie-breaking rule that for equal revenues, A prefers to tax the domestic tax base, for $t_A < \bar{t}$ to be the best response it would have to be that $TR_A^* > d\bar{t}\Pi$, which boils down to $\alpha < \alpha^{thr}$. Thus, $t_A = t_B = \bar{t}$ is an equilibrium under $\alpha \geq \alpha^{thr}$.

For uniqueness, consider a situation with an arbitrary $t_B \in [0, \bar{t})$. Again, choosing $t_A = \bar{t}$ yields a certain tax revenue of $d\bar{t}\Pi$. Further, any $t_A \in [t_B, \bar{t})$ implies that A only taxes the domestic tax base at a rate strictly lower than $\bar{t}$ and thus, is pay-off dominated by $\bar{t}$. Finally, for any $t_A \in [0, t_B]$ A certainly becomes the Stackelberg-leader in a competition with C. However, due to strict concavity, tax revenues will be weakly lower than TR_A^* and thus, any such choice is dominated by $t_A = \bar{t}$ as well. Thus, it could be shown that $t_A = \bar{t}$ is the dominant strategy for A and thus, $t_A = t_B = \bar{t}$ is the unique equilibrium under $\alpha \geq \alpha^{thr}$.

(ii)

Without loss of generality, consider a situation with $t_A > t_B$. A will only host the domestic tax base and thus, in equilibrium it will charge the domestic Laffer rate $t_A = \bar{t}$. Further, B will certainly compete with C in a Stackelberg game and thus, in equilibrium it will choose the international Laffer rate $t_B = t_B^* < \bar{t}$. Consider a deviation for A. Given continuity of Stackelberg revenues around t_B^*, A can always marginally undercut t_B^* to earn its maximum level of revenues from the competition with C. Under $\alpha < \alpha^{thr}$ Stackelberg revenues exceed revenues from domestic taxation and thus, $t_A = \bar{t}$ is not the best response to $t_B = t_B^*$. Since it has been shown that these tax rates are the only candidates for an equilibrium with $t_A > t_B$, there does not exist an equilibrium with $t_A > t_B$. Next, consider the case of $t_A = t_B = t$. Generally assume that in such a situation, the mass of mobile firms not paying taxes in C is divided between A and B at rates μ_A, μ_B, with $\mu_A + \mu_B = 1$. Without loss of generality assume that $\mu_A \leq \mu_B$. The condition implies that $\mu_A \leq 0.5$ and guarantees that $TR^A(\mu_A, t) \leq TR^B(\mu_B, t)$. The tax revenues at t can be written as $TR_A(t) = t\,(d + \mu_A m(t))$, where $m(t)$ gives the overall mass of mobile firms not paying taxes in C. Instead, consider a deviation to $t_A = t - \epsilon$. The tax revenues in this case read $TR_A(t-\epsilon) = (t-\epsilon)\,(d + m(t-\epsilon))$. For $t - \epsilon$ to be an optimal deviation it has to be that $TR_A(t-\epsilon) \geq TR_A(t)$, which boils down to $\epsilon \leq \frac{m(t-\epsilon) - \mu_A m(t)}{d + m(t-\epsilon)}$. For any $t \in [0, \bar{t}]$, $m(t-\epsilon) \geq m(t)$ holds in the model and thus, there always exists an ϵ arbitrarily close to zero such that A will find it optimal to marginally undercut t, if $m(t-\epsilon) > 0$. See that a necessary condition for the existence of an equal tax rate equilibrium is $TR^A(\mu_A, t) \geq d\bar{t}\Pi$, otherwise A would deviate to $t_A = \bar{t}$. Accordingly, in any equal tax equilibrium with $t < \bar{t}$, $m(t)$ has to be strictly larger than zero and thus, undercutting t marginally is always an optimal deviation for any $t < \bar{t}$. Instead, for $t = \bar{t}$ it is not guaranteed that $m(t-\epsilon) > 0$ for an ϵ arbitrarily close to t. However, at $\bar{t}$ A could choose ϵ such that $\bar{t} - \epsilon = t_A^*$ and by the condition on $\alpha < \alpha^{thr}$ increase tax revenues. Thus, even in the case of $t = \bar{t}$, there exists a (non-marginal) ϵ such that $t_A = t$ is not the optimal response. So, it could be shown that there does not exist an equilibrium in pure strategies for $\alpha < \alpha^{thr}$.

(iii)

In a mixed equilibrium with support Ω and density function $g(t_i)$, the expected tax revenues for any $t_A \in \Omega$ are given by:

$$\mathbb{E}\left(TR_A(t_A, t_B)\right) = \int_{t_B \in \Omega} g(t_B) TR_A(t_A, t_B) dt_B \tag{A.5}$$

The construction of the mixed equilibrium proceeds in two steps. First, the equilibrium strategy supports, Ω, and the probability of playing $\bar{t}$, $Pr(\bar{t})$, are determined. Second, the mixing probabilities for the convex interval $[t', t'']$ are derived. Without loss of generality, I will focus on the strategies of A.

Proposition 2 has shown that Stackelberg tax revenues for $t_A > 2\frac{rF}{\Pi}\frac{\Delta}{\alpha}$ are given by $at_A\Pi$. Thus, independent of t_B, any $t_A \in (2\frac{rF}{\Pi}\frac{\Delta}{\alpha}, \bar{t})$ is strictly dominated by $t_A = \bar{t}$. Stackelberg tax revenues are differentiable and concave on $t_A \in [0, 2\frac{rF}{\Pi}\frac{\Delta}{\alpha}] \equiv S$ and $t_A^* \in S$. The revenues at the boundaries of S are 0 for $t_A = 0$ and, by continuity, $2\frac{rF}{\Pi}\frac{\Delta}{\alpha}d\Pi$ at the upper bound. Under the maximum degree of differentiation Δ^{max}, it can be shown that $2\frac{rF}{\Pi}\frac{\Delta}{\alpha}d\Pi \leq \bar{t}d\Pi$. Further, under $\alpha < \alpha^{thr}$, Stackelberg revenues at t_A^* exceed revenues from domestic taxation. Then, by differentiability and concavity of tax revenues on S, there exist two tax rates $t_{low}, t_{up} \in S$, with $t_{low} \leq t_A^* \leq t_{up}$, such that for any $t \in S \backslash [t_{low}, t_{up}] \equiv S'$, $TR_A(t) < a\bar{t}\Pi$. Thus, independent of t_B, any $t \in S'$ is dominated by $\bar{t}$. So, in a first step, the candidates for an equilibrium support have been restricted to $\Omega^{cd} = \left\{[t_{low}, t_{up}]; \bar{t}\right\}$.

In a next step, I will show that in any equilibrium there does not exist a mass point on $[t_{low}, t_{up}]$. Assume there exists a $t_A = t_B \in [t_{low}, t_{up}]$, which is played with positive probability $\beta > 0$. Further, B plays a tax rate larger than t with probability γ. Applying the structure from (ii), A will find it optimal to replace the probability mass on t with (additional) probability mass on $t - \epsilon$, if $\epsilon \leq \frac{(\beta+\gamma)m(t-\epsilon)-(\mu_A\beta+\gamma)m(t)}{d+(\beta+\gamma)m(t-\epsilon)}t$. Given that by construction of the interval, for every $t \in [t_{low}, t_{up}]$ it holds that $m(t-\epsilon) \downarrow m(t) > 0$ for $\epsilon \to 0$, there always exists an ϵ sufficiently close to zero such that A would substitute the probability mass on t with (additional) probability mass on $t - \epsilon$.

In a next step, I will narrow down Ω^{cd} to the actual equilibrium support Ω. Initially, I will determine the upper bound of the convex interval, t''. It was shown before that at t'' there does not exist a mass point in equilibrium. Further, at t'' the expected tax revenues have to equal the revenues from domestic taxation. Thus, the first condition on t'' is given by:

$$\left(1 - Pr(\bar{t})\right) dt''\Pi + Pr(\bar{t})t'' \left(d + m(t'')\right) \Pi = d\bar{t}\Pi \tag{A.6}$$

Here, $Pr(\bar{t})$ represents the probability mass at $t_B = \bar{t}$. The condition will never be fulfilled for $t'' = t_{up}$ and thus $t'' < t_{up}$. Given that $(t'', \bar{t})$ was shown not to be part of the equilibrium support, marginally increasing the tax rate at t'' will not change the Probability of underbidding B. On the other hand, increasing t at t'' will increase revenues in the case of domestic taxation and affects Stackelberg revenues as well. Thus, t'' is set such that a marginal increase does not constitute an optimal deviation. This second condition is given by:

$$\left(1 - Pr(\bar{t})\right) d\Pi + Pr(\bar{t}) \left(d + m(t'') + t''\frac{\partial m(t'')}{\partial t}\right) \Pi = 0 \tag{A.7}$$

By the First Order Condition for a local optimum, the in-bracket part of the second term will be equal to zero at $t'' = t_A^*$. Further, by concavity, this implies that $t'' \geq t_A^*$. Finally using continuity, it can be stated that $t'' \in (t_A^*, t_{up})$. See that (A.6) and (A.7) constitute a system of two (nonlinear) equations in two unknowns, $Pr(\bar{t})$ and t''. Applying the functional forms to (A.7) and solving for t'' yields:

$$t'' = \frac{rF}{\Pi}\frac{\Delta}{\alpha}\left(1 + \frac{d}{Pr(\bar{t})m}\right)$$

Using the result in (A.6), the positive root of the quadratic expression is given by:

$$Pr\left(\bar{t}\right) = \frac{d}{m}\left(\frac{\Pi - \frac{r}{\theta}F}{\frac{r}{\theta}F}\frac{\alpha}{\left(\frac{\bar{\theta}}{\underline{\theta}} - 1\right)} - 1\right)\left[1 + \sqrt[2]{1 - \frac{1}{\left(\frac{\Pi - \frac{r}{\theta}F}{\frac{r}{\theta}F}\frac{\alpha}{\left(\frac{\bar{\theta}}{\underline{\theta}} - 1\right)} - 1\right)^2}}\right]$$

For any set of exogenous parameters and $\alpha < \alpha^{thr}$, in equilibrium there exists a unique $Pr(\bar{t}) > 0$. Using this result, there exists a unique t'' as presented above.

In a next step, I will show that in any equilibrium support it has to be that $t' = t_{low}$. Suppose not. Initially see that by the construction of t_{low}, it has to be that $t' \geq t_{low}$. So, without loss of generality, consider the best response of A to any mixing strategy of B on $\Omega_B = \left\{[t', t'']; \bar{t}\right\}$ with $t' > t_{low}$. Given that $\bar{t}$ is contained in the equilibrium support, any mixing equilibrium yields expected revenues of $a\bar{t}\Pi$. On the other hand, charging a $t_A = t' - \epsilon$ with ϵ sufficiently close to zero, yields the Stackelberg revenues with certainty. Given that $t' > t_{lower}$, by the construction of t_{lower}, Stackelberg revenues exceed revenues from domestic taxation and thus, the best response to Ω_B is given by $t_A = t' - \epsilon$. So, I have shown that $t' \neq t_{low}$ can never be sustained in a mixed equilibrium support and thus, in any Ω^* it has to be that $t' = t_{low}$.

It is left to derive the equilibrium mixing probabilities for the convex part of Ω. For any $t_A \in [t', t'']$, (A.5) is given by:

$$G_B(t_B > t_A)TR_A\left(t_A|t_B > t_A\right) + G_B(t_B \leq t_A)TR_A\left(t_A|t_B \leq t_A\right) = a\bar{t}\Pi$$

It was previously shown that on the convex interval, there does not exist a mass point in any equilibrium and thus for $t_A \in [t', t'']$, $Pr(t_A = t_B) = 0$. Thus, denote by $G_B(t_B \leq t_A)$ the cumulative probability that B chooses a tax rate weakly lower than t_A. The condition can be reformulated such that:

$$G_B\left(t_B \leq t_A\right) = \frac{TR_A\left(t_A|t_B > t_A\right) - d\bar{t}\Pi}{TR_A\left(t_A|t_B > t_A\right) - TR_A\left(t_A|t_A > t_B\right)}$$

Now, see that by symmetry of the supports and tax revenues it has to be that $G_B\left(t_B \leq t_A\right) = G(t_A)$, with $G(t_A)$ being the cumulative distribution function of t_A. Finally, using the definitions of t' and t'' it can be easily verified that $G(t_A = t') = 0$ and $G(t_A = t'') = 1 - Pr(\bar{t})$.

Proof of Proposition 4

Before the start of the Proof, consider the distribution of the minimum tax rate of high productive countries in a mixed strategy equilibrium, t_{min}. On $[t', t'']$, tax rates of the high productive countries are random variables distributed with respect to two identical and independent cumulative distributions G(t) and density function g(t). Accordingly, the density of the minimum is given by $\Pr(t = t_{min}) = 2g(t)(1 - G(t))$ and the cumulative distribution by $\Pr(t \leq t_{min}) = 2G(t)(1 - 0.5G(t))$.

(i) I want to show that, given that $\alpha^{thr} > 1$, for every $\alpha \in \left[1, \alpha^{thr}\right]$, the marginal inefficiency of an uplift is always positive. The marginal inefficiency is given by:

$$\frac{\partial IE}{\partial \alpha} = \underbrace{m\int_{t'}^{t''} \frac{\partial f^{thr}(t)}{\partial \alpha} \text{ACC}^{f^{thr}(t)} \Pr(t_{min} = t) dt}_{\text{extensive firm}}$$

$$+ \underbrace{\Pr\left(\bar{t}\right)^2 m \int_0^1 \frac{\partial \text{ACC}^f}{\partial \alpha} df + m \int_{t'}^{t''} \left[\int_0^{f^{thr}(t)} \frac{\partial \text{ACC}^f}{\partial \alpha} df\right] \Pr\left(t = t_{min}\right) dt}_{\text{intensive firm}}$$

$$+ \underbrace{2\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha} \Pr\left(\bar{t}\right) m \left[\int_0^1 \text{ACC}^f df - \int_0^{f^{thr}(t'')} \text{ACC}^f df\right] - m\int_{t'}^{t''} \left\{\frac{\partial \left[\int_0^{f^{thr}(t)} \text{ACC}^f df\right]}{\partial t} \frac{\partial \Pr\left(t_{min} \leq t\right)}{\partial \alpha}\right\} dt}_{\text{strategic leader}}$$

The remainder of the proof proceeds as follows: First, it will be shown that the sum of extensive firm effect and the second term of the intensive firm effect is unambiguously positive. Second, it will be shown that the sum of the first term of the intensive firm effect and the first term of the strategic leader effect is unambiguously positive. If that holds, the marginal inefficiency is unambiguously positive, since the second term of the strategic leader effect is unambiguously positive, due to $\frac{\partial \left[\int_0^{f^{thr}(t)} ACC^f df\right]}{\partial t} > 0$ and $\frac{\partial Pr(t_{min} \leq t)}{\partial \alpha} < 0$.

In the first step, it thus has to be shown that:

$$m\int_{t'}^{t''} \frac{\partial f^{thr}(t)}{\partial \alpha} \text{ACC}^{f^{thr}(t)} \Pr\left(t_{min} = t\right) dt + m\int_{t'}^{t''} \left[\int_0^{f^{thr}(t)} \frac{\partial \text{ACC}^f}{\partial \alpha} df\right] \Pr\left(t_{min} = t\right) dt \geq 0$$

Merging the two outer intergrals, solving the inner integral and plugging in the respective closed forms, the condition reformulates to:

$$m\int_{t'}^{t''} \left[\frac{1}{\Delta} 0.5t \frac{\Pi}{rF} 0.5t\Pi - \frac{\Delta}{\alpha^2} rF 0.5 \left(\frac{\alpha}{\Delta} 0.5t \frac{\Pi}{rF}\right)^2\right] \Pr\left(t_{min} = t\right) dt \geq 0$$

Simplifying the terms gives:

$$0.5m\int_{t'}^{t''} \frac{1}{\Delta} \frac{\Pi}{rF} \left(0.5t\right)^2 \Pi \Pr\left(t_{min} = t\right) dt \geq 0$$

The condition unambiguously holds true. So, the first part of the proof is completed, the sum of the extensive firm effect and the second term of the intensive firm effect is unambiguously positive.

I proceed with the second part of the proof, which seeks to show that the sum of the first term of the intensive firm effect and the first term of the strategic leader effect is unambiguously positive. It thus has to be shown that:

$$m\Pr\left(\bar{t}\right)^2 \int_0^1 \frac{\partial \text{ACC}^f}{\partial \alpha} df + 2\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha} \Pr\left(\bar{t}\right) m \left[\int_0^1 \text{ACC}^f df - \int_0^{f^{thr}(t'')} \text{ACC}^f df\right] \geq 0$$

Using the closed forms for ACC, the condition reformulates to:

$$-m\Pr\left(\bar{t}\right)^2 0.5\frac{\Delta}{\alpha^2} rF + m2\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha} \Pr\left(\bar{t}\right) \left[0.5\frac{\Delta}{\alpha} rF - 0.5 f^{thr}\left(t''\right)^2 \frac{\Delta}{\alpha} rF\right] \geq 0$$

Simplifying the condition and using the closed forms of f^{thr} and t'' gives:

$$2\frac{\frac{\partial \Pr(\bar{t})}{\partial \alpha}}{\Pr\left(\bar{t}\right)}\alpha\left[1-\left(0.5\left(1+\frac{d}{\Pr\left(\bar{t}\right)m}\right)\right)^2\right]\geq 1$$

From here on the closed form on $\Pr\left(\bar{t}\right)$ is used to simplify the expression further. For the ease of exposition, define $x \equiv \frac{\bar{t}}{1-\bar{t}}\frac{\alpha}{\frac{\bar{\theta}}{\underline{\theta}}-1}$. See that under $\alpha^{thr} > 1$, it holds that $x > 2$. Then, the expression reads:

$$2\frac{x}{x-1}\left[1+\frac{1}{\sqrt{x(x-2)}\left(x-1+\sqrt{x(x-2)}\right)}\right]\left[1-\left(0.5\left(1+\frac{1}{\left(x-1+\sqrt{x(x-2)}\right)}\right)\right)^2\right]\geq 1$$

It is easy to show from here, that the LHS of the condition is monotonically decreasing in x and so ultimately, decreasing in α. Accordingly, it is sufficient to show that the condition holds for the largest feasible α, in the model α^{thr}. In particular, the threshold is given by $\alpha^{thr} = 0.5\frac{(d+m)^2}{dm}\frac{1-\bar{t}}{\bar{t}}\bar{\theta}\Delta$, such that $x\left(\alpha^{thr}\right) = 0.5\frac{(d+m)^2}{dm}$. Using this expression, the condition reformulates to:

$$2\frac{(d+m)^2}{d^2+m^2}\frac{m^2}{m^2-d^2}\frac{m^2-0.25\left(m+d\right)^2}{m^2}\geq 1$$

The condition reformulates to:

$$0.75m^2d + 0.25d^3 + 0.25m\left(m^2-d^2\right)\geq 0$$

Given the assumption of $m \geq d$, the condition unambiguously holds true. So, it has been shown that the sum of the first term of the intensive firm effect and the first term of the strategic leader effect is unambiguously positive. In total, it has been now shown that the marginal inefficiency of an uplift is unambiguously positive on $\alpha \in \left[1, \alpha^{thr}\right]$.

(ii) Now, show that the marginal inefficiency is unambiguously negative for $\alpha > \alpha^{thr}$. Given that high productive countries selected out of competition for $\alpha > \alpha^{thr}$ already, the marginal inefficiency in such case collapses to:

$$\frac{\partial \mathrm{IE}}{\partial \alpha} = m\int_0^1 \frac{\partial \mathrm{ACC}^f}{\partial \alpha}df$$

Using the closed form of ACC, it is easy to show that the marginal inefficiency is unambiguously negative in that case.

Proof of Lemma 3

The perfectly mobile firm is indexed by $f = 0$. Given that the derivative of the firm-specific marginal additional capital costs is given by $\frac{\partial \mathrm{ACC}^f}{\partial \alpha} = -f\frac{\Delta}{\alpha^2}rF$, for the perfectly mobile firm, the total effect is equal to the strategic effect and thus, profits unambiguously decrease in the level of uplift.

Proof of Lemma 4

The least mobile firm will never invest in the low productive country, as long as at least one high productive country is actively competing for mobile tax revenues, $t_{min} < \bar{t}$. Consequentially, the marginal profits are given by:

$$\frac{\partial \mathrm{MoP}^1}{\partial \alpha} = -\Pr\left(\bar{t}\right)^2\frac{\partial t^m}{\partial \alpha}\Pi - \Pr\left(\bar{t}\right)^2\frac{\partial \mathrm{ACC}^f}{\partial \alpha} + \text{strategic effect}$$

Where the last part captures the strategic effect, which is unambiguously negative. Thus, a sufficient condition for negativity is $-\frac{\partial t^m}{\partial \alpha}\Pi < \frac{\partial \mathrm{ACC}^1}{\partial \alpha}$, which unambiguously holds true. Thus, the marginal profits of the least mobile firm unambiguously decrease in the level of uplift.

Proof of Proposition 5

The proof shows that it will be a firm with an intermediate degree that maximizes marginal profits from a lift in the limitation on profit shifting.

Generally, marginal profits for a mobile firm with a degree of mobility indexed by f are given by:

$$\frac{\partial \mathrm{MoP}^f}{\partial \alpha} = -\underbrace{\Pr(\bar{t})^2 \frac{\partial t^m}{\partial \alpha}\Pi}_{\text{strategic leader}} - \underbrace{\left[\Pr(\bar{t})^2 \frac{\partial \mathrm{ACC}^f}{\partial \alpha} + \int_{t'}^{t''} \left\{\mathbb{I}_{f^{thr}(t)\geq f} \frac{\partial \mathrm{ACC}^f}{\partial \alpha}\right\} \Pr(t_{min} = t)\right]}_{\text{intensive firm}}$$
$$\underbrace{+ 2\frac{\partial \Pr(\bar{t})}{\partial \alpha} \Pr(\bar{t}) \left[\Pi_L^f(t^m) - \min\left\{\Pi_L^f\left(t_n(t'')\right); \Pi_H(t'')\right\}\right] - \int_{t'}^{t''} \frac{\partial \min\left\{\Pi_L^f\left(t_n(t)\right); \Pi_H(t)\right\}}{\partial t} \frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha}}_{\text{strategic leader}}$$

Marginal profits are generally continuous on $f \in [0, 1]$ and piecewise differentiable. The continuum of mobile firms can be split up in three parts. First firms with a high degree of mobility, a small f, decide to invest in the low productive country for every $t \in [t', t'']$. In particular, there exists a f^{low}, such that all firms with $f \leq f^{low}$ invest in the low productive country for every t on the convex interval. Second firms with a low degree of mobility, a larger f, decide not to invest in the low productive country for every $t \in [t', t'']$. In particular, there exists a f^{high}, such that all firms with $f \geq f^{high}$ do not invest in the low productive country for every t on the convex interval. Finally for every firm with a degree of mobility $f \in \left(f^{low}, f^{high}\right)$ there exists a $t(f) \in (t', t'')$, such that for $t \geq t(f)$ the firm does invest in the low productive country and for $t < t(f)$ it does not invest in the low productive country.

I will proceed showing that marginal profits are increasing in f on $f \in \left[0, f^{low}\right]$ and decreasing in f on $f \in \left[f^{low}, 1\right]$. This, taking into account continuity, implies that marginal profits are the highest for an intermediate degree of mobility, f^{low}.

Start with $f \leq f^{low}$. In this case the derivative simplifies to:

$$\left.\frac{\partial \mathrm{MoP}^f}{\partial \alpha}\right|_{f \leq f^{low}} = -\Pr\left(\bar{t}\right)^2 \frac{\partial t^m}{\partial \alpha}\Pi - \frac{\partial \mathrm{ACC}^f}{\partial \alpha}$$
$$+ 2\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha} \Pr\left(\bar{t}\right) \left[\Pi_L^f\left(t^m\right) - \Pi_L^f(t_n\left(t''\right)\right] - \int_{t'}^{t''} \frac{\partial \Pi_L^f\left(t_n(t)\right)}{\partial t} \frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha}$$

The derivative with respect to f reads:

$$\left.\frac{\partial^2 \mathrm{MoP}^f}{\partial \alpha \partial f}\right|_{f \leq f^{low}} = -\frac{\partial^2 \mathrm{ACC}^f}{\partial \alpha \partial f} > 0$$

So, it could be shown that marginal profits increase in f on $f \in \left[0, f^{low}\right]$.

In the next step, discuss the case of $f \geq f^{high}$. In this case the derivative simplifies to:

$$\left.\frac{\partial \text{MoP}^f}{\partial \alpha}\right|_{f \geq f^{high}} = -\Pr\left(\bar{t}\right)^2 \frac{\partial t^m}{\partial \alpha}\Pi - \Pr\left(\bar{t}\right)^2 \frac{\partial \text{ACC}^f}{\partial \alpha} + 2\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha}\Pr\left(\bar{t}\right)\left[\Pi_L^f\left(t^m\right) - \Pi_H^f\left(t''\right)\right] - \int_{t'}^{t''} \frac{\partial \Pi_H^f\left(t\right)}{\partial t}\frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha}$$

The derivative with respect to f reads:

$$\left.\frac{\partial^2 \text{MoP}^f}{\partial \alpha \partial f}\right|_{f \geq f^{high}} = -\Pr\left(\bar{t}\right)^2 \frac{\partial^2 \text{ACC}^f}{\partial \alpha \partial f} + 2\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha}\Pr\left(\bar{t}\right)\frac{\partial \Pi_L^f\left(t_M\left(t''\right)\right)}{\partial f}$$

Using the functional forms of $\Pr\left(\bar{t}\right)$, ACC^f and Π_L^f it can be easily shown that unambiguously $\left.\frac{\partial^2 MoP^f}{\partial \alpha \partial f}\right|_{f \geq f^{high}} < 0$. Thus marginal profits decrease in f for $f \in \left[f^{high}, 1\right]$.

Finally discuss the case of $f \in \left(f^{low}, f^{high}\right)$. In this case, using the definition of $t(f)$, the derivative simplifies to:

$$\left.\frac{\partial \text{MoP}^f}{\partial \alpha}\right|_{f \in \left(f^{low}, f^{high}\right)} = -\left[1 - \Pr\left(t_{min} \leq t(f)\right)\right]\frac{\partial \text{ACC}^f}{\partial \alpha} + 2\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha}\Pr\left(\bar{t}\right)\left[\Pi_L^f\left(t_m\left(t''\right)\right) - \Pi_L^f\left(t_N\left(t''\right)\right)\right] - \Pr\left(\bar{t}\right)^2\frac{\partial t^m}{\partial \alpha}\Pi - \int_{t'}^{t(f)} \frac{\partial \Pi_H^f\left(t\right)}{\partial t}\frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha} - \int_{t(f)}^{t''} \frac{\partial \Pi_L^f\left(t_n\left(t\right)\right)}{\partial t}\frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha}$$

The derivative with respect to f reads:

$$\left.\frac{\partial^2 \text{MoP}^f}{\partial \alpha \partial f}\right|_{f \in \left(f^{low}, f^{high}\right)} = -\left[1 - \Pr\left(t_{min} \leq t(f)\right)\right]\frac{\partial^2 \text{ACC}^f}{\partial \alpha \partial f} + \frac{\partial t(f)}{\partial f}\Pr\left(t_{min} = t(f)\right)\frac{\partial \text{ACC}^f}{\partial \alpha} - \frac{\partial t(f)}{\partial f}\frac{\partial \Pi_H\left(t(f)\right)}{\partial t}\frac{\partial \Pr\left(t_{min} \leq t(f)\right)}{\partial \alpha} + \frac{\partial t(f)}{\partial f}\frac{\partial \Pi_L^f\left(t_n(t(f))\right)}{\partial t}\frac{\partial \Pr\left(t_{min} \leq t(f)\right)}{\partial \alpha}$$

Using the functional forms it can easily be shown that $\left.\frac{\partial^2 \text{MoP}^f}{\partial \alpha \partial f}\right|_{f \in \left(f^{low}, f^{high}\right)} < 0$ holds for $\bar{t} \geq 2\Delta\frac{rF}{\Pi}$, which was one of the model's assumptions. Thus marginal profits decrease in f for $f \in \left(f^{low}, f^{high}\right)$.

In sum, it was shown that marginal profits increase in f up to $f = f^{low}$ and decrease thereafter. Thus, by continuity, it will be an intermediate firm with $f = f^{low}$ that increases the most from a lift in the limitation on Profit Shifting at $\alpha = 1$.

Proof of Proposition 6

If a high productive country does not compete for mobile firms, the national income is given by:

$$\text{I}^{\text{aut}} = d\left(\Pi - \frac{r}{\bar{\theta}}F\right)$$

The expression is independent of the limitation on profit shifting. Further, the marginal revenues of a high productive country, competing for mobile firms, is given by:

$$\frac{\partial \text{I}^{\text{mix}}}{\partial \alpha} = d\Pi\left[-\left(\bar{t} - t''\right)\frac{\partial \Pr\left(\bar{t}\right)}{\partial \alpha} + \int_{t'}^{t''} \frac{\partial G(t)}{\partial \alpha}\right]$$

Using the functional forms, it is straightforward to show that $\bar{t} > t''$, $\frac{\partial \Pr(\bar{t})}{\partial \alpha} > 0$ and $\frac{\partial G(t)}{\partial \alpha} < 0 \ \forall t \in [t', t'']$. Therefore it holds that $\frac{\partial I^{mix}}{\partial \alpha}$. In sum, it has been shown that national income of a high productive country decrease in the limitation on profit shifting on $\alpha \in \left[1, \alpha^{thr}\right]$ and is independent on the limitation on profit shifting on $\alpha > \alpha^{thr}$.

Proof of Proposition 7

I will derive the derivative of the low productive country's tax revenues with respect to α, starting from expression (5). Plugging in and applying integration by parts to the integral, tax revenues reformulate:

$$\begin{aligned}\text{TR} &= mt^m \Pi \Pr\left(\bar{t}\right)^2 \\ &+ m\Pi \left[\left[t^n(t) f^{thr}(t) 2G(t)\left(1-0.5G(t)\right)\right]_{t'}^{t''} - \int_{t'}^{t''} \left(\frac{\partial t^n(t)}{\partial t} f^{thr}(t) + t^n(t) \frac{\partial f^{thr}(t)}{\partial t} \right) 2G(t)(1-0.5G(t)) \right]\end{aligned} \tag{A.8}$$

From here, use the functional forms of $t^n(t)$ and $f^{thr}(t)$ as well as the boundary properties on the distribution $G(t') = 0$ and $G(t'') = 1 - \Pr(\bar{t})$. The revenues reformulate once more:

$$\begin{aligned}\text{TR} &= mt^m \Pi \Pr(\bar{t})^2 \\ &+ m\Pi \left[t^n(t'') f^{thr}(t'') \left(1 - \Pr(\bar{t})^2\right) - \int_{t'}^{t''} f^{thr}(t) 2G(t)(1-0.5G(t)) \right]\end{aligned} \tag{A.9}$$

Now, pull the derivative using the Leibniz integral rule to obtain:

$$\begin{aligned}\frac{\partial \text{TR}}{\partial \alpha} &= m\frac{\partial t^m}{\partial \alpha} \Pi \Pr(\bar{t})^2 + mt^m \Pi 2 \frac{\partial \Pr(\bar{t})}{\partial \alpha} \Pr(\bar{t}) \\ &+ m\Pi \left[0.5t'' \frac{\partial t''}{\partial \alpha} \frac{\alpha}{\Delta} \frac{\Pi}{rF} \left(1 - \Pr(\bar{t})^2\right) + (0.5t'')^2 \frac{1}{\Delta} \frac{\Pi}{rF} \left(1 - \Pr(\bar{t})^2\right) - t^n(t'') f^{thr}(t'') 2 \frac{\partial \Pr(\bar{t})}{\partial \alpha} \Pr(\bar{t}) \right] \\ &+ m\Pi \left[-\frac{\partial t''}{\partial \alpha} f^{thr}(t'') 2G(t'')(1-0.5G(t'')) - \int_{t'}^{t''} \left[\frac{\partial f^{thr}}{\partial \alpha} 2G(t)\left(1-0.5G(t)\right) + f^{thr}(t) 2 \frac{\partial G(t)}{\partial \alpha} (1-G(t)) \right] \right]\end{aligned} \tag{A.10}$$

Using the functional forms and boundary properties once more, the derivative can be simplified to:

$$\begin{aligned}\frac{\partial \text{TR}}{\partial \alpha} &= m\frac{\partial t^m}{\partial \alpha} \Pi \Pr(\bar{t})^2 \\ &+ m\Pi \left[2\frac{\partial \Pr(\bar{t})}{\partial \alpha} \Pr(\bar{t}) \left(t^m - t^n(t'') f^{thr}(t'')\right) - \int_{t'}^{t''} f^{thr}(t) 2 \frac{\partial G(t)}{\partial \alpha} (1-G(t)) \right] \\ &+ m\Pi \left[t^n(t'') \frac{\partial f^{thr}(t'')}{\partial \alpha} \left(1 - \Pr(\bar{t})^2\right) - \int_{t'}^{t''} \frac{\partial f^{thr}}{\partial \alpha} 2G(t)\left(1-0.5G(t)\right) \right]\end{aligned} \tag{A.11}$$

Using integration by parts, the last in-bracket term reformulates to $\int_{t'}^{t''} t^n(t) \frac{\partial f^{thr}(t)}{\partial \alpha} 2g(t)(1-G(t))$. Thus, (A.11) is equivalent to (13). The first and third term being positive follows directly from the functional forms of t^m and $f^{thr}(t)$. Positivity of the second term follows from $\frac{\partial \Pr(\bar{t})}{\partial \alpha} > 0 > \frac{\partial G(t)}{\partial \alpha}$ and $\bar{t} \geq 2\frac{rF}{\Pi}\frac{\Delta}{\alpha} \geq t''$. Finally, see that the third term captures the mechanic impact of a lift in the limitation, whereas the first two terms capture strategic effect. Given that each individual term is positive, it could be shown that strategic effects reinforce the mechanic effect to increase tax revenues of low productive countries.

Proof of Proposition 8

Regarding (i), the mechanic effect is given by:

$$\text{ME} = -\left[\Pr(\bar{t})^2 m \int_0^1 \frac{\partial \text{ACC}^f}{\partial \alpha} df - \int_{t'}^{t''} \left\{\left[\int_0^{f^{thr}(t)} \frac{\partial \text{ACC}^f}{\partial \alpha} df\right] \Pr(t_{min} = t)\right\} dt\right]$$

It has been shown that the marginal additional capital costs are given by $\frac{\partial \text{ACC}^f}{\partial \alpha} = -f\frac{\Delta}{\alpha^2} rF$. Thus, the mechanic effect is unambiguously positive. Regarding (ii), the strategic effect is given by:

$$\text{SE} = -\Pr(\bar{t})^2 m \frac{\partial t^m}{\partial \alpha}\Pi + 2\frac{\partial \Pr(\bar{t})}{\partial \alpha}\Pr(\bar{t})\left[\text{MoP}^m - \text{MoP}^n(t'')\right] - m\int_{t'}^{t''} \left\{\frac{\partial \text{MoP}^n(t)}{\partial t}\frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha}\right\} dt$$

The first term is negative, since $\frac{\partial t^m}{\partial \alpha} > 0$. The second term is negative, since $\frac{\partial \Pr(\bar{t})}{\partial \alpha} > 0$ and $\text{MoP}^m < \text{MoP}^n(t'')$. Finally, the third term is negative, given that $\frac{\partial \text{MoP}^n(t)}{\partial t}, \frac{\partial \Pr(t_{min} \leq t)}{\partial \alpha} > 0 \;\; \forall t \in [t', t'']$. Thus, the strategic effect is unambiguously negative. The proof is left to determine the total effect, TE = ME + SE. Using the functional forms and rearranging, the total effect derives as:

$$\begin{aligned}\text{TE} = &- m\Pi\left[2\frac{\partial \Pr(\bar{t})}{\partial \alpha}\Pr(\bar{t})\left(\left(\bar{t} - \frac{rF}{\Pi}\frac{\Delta}{\alpha} - \frac{(0.5t'')^2}{\frac{\Delta}{\alpha}\frac{rF}{\Pi}} - \frac{d}{m}\frac{(\bar{t} - t'')}{\Pr(\bar{t})}\right) + 0.5\left(\frac{\Delta}{\alpha}\frac{rF}{\Pi} - \frac{(0.5t'')^2}{\frac{\Delta}{\alpha}\frac{rF}{\Pi}}\right)\right)\right.\\ &+ 0.5\Pr(\bar{t})^2\frac{rF}{\Pi}\frac{\Delta}{\alpha^2} + \frac{3}{2}\frac{1}{\Delta}\frac{\Pi}{rF}\left[0.25\left(1 - \Pr(\bar{t})^2\right)(t'')^2 - \int_{t\in[t',t'']} tF(t)(1 - 0.5F(t))\right]\\ &\left.- \int_{t\in[t',t'']} \frac{\partial F(t)}{\partial \alpha}\left[\frac{3}{2}\frac{\alpha}{\Delta}\frac{\Pi}{rF}t(1 - F(t)) - 2\frac{d}{m}\right]\right]\end{aligned}$$

Thus, for the Total Effect to be negative, the sum of terms in the most outer bracket has to be positive. While it is straightforward to see that the second and third term are unambiguously positive, the determination of the sign for the first and fourth term is a bit more involved. In particular, given that the second term is always strictly positive, to show that the in-bracket term is overall strictly positive, it suffices to show that the first and fourth term are non-negative. Consider the first term. Given that $\frac{\partial \Pr(\bar{t})}{\partial \alpha} > 0$, for the first term to be nonnegative it has to hold that:

$$\left(\bar{t} - \frac{rF}{\Pi}\frac{\Delta}{\alpha} - \frac{(0.5t'')^2}{\frac{\Delta}{\alpha}\frac{rF}{\Pi}} - \frac{d}{m}\frac{(\bar{t} - t'')}{\Pr(\bar{t})}\right) + 0.5\left(\frac{\Delta}{\alpha}\frac{rF}{\Pi} - \frac{(0.5t'')^2}{\frac{\Delta}{\alpha}\frac{rF}{\Pi}}\right) \geq 0$$

For the second in-bracket part to be nonnegative it has to hold that $t'' \leq 2\frac{\Delta}{\alpha}\frac{rF}{\Pi}$, which unambiguously holds true on $\alpha \in \left[1, \alpha^{thr}\right]$. Regarding the first term, after some reformulation, non-negativity requires:

$$\bar{t}\left(1 - \frac{d}{\Pr(\bar{t})\, m}\right) + t''\left(\frac{d}{m\Pr(\bar{t})} - \frac{0.25t''}{\frac{\Delta}{\alpha}\frac{rF}{\Pi}}\right) \geq \frac{rF}{\Pi}\frac{\Delta}{\alpha}$$

Given the closed form of t'', $t'' = \frac{rF}{\Pi}\frac{\Delta}{\alpha}\left(1 + \frac{d}{\Pr(\bar{t})m}\right)$, the condition yields:

$$1 \geq \frac{d}{m\Pr(\bar{t})}\left(2 - \frac{d}{m\Pr(\bar{t})}\right)$$

It has been shown that on $\alpha \in \left[1, \alpha^{thr}\right]$, it holds that $\Pr(\bar{t}) \in \left[\frac{d}{m}, 1\right]$ and so, the condition unambiguously holds.

Thus, to determine the overall sign of the derivative, the determination of the sign of the fourth term is left. It has been shown that $\frac{\partial F(t)}{\partial \alpha} < 0$. Thus, non-negativity of the fourth term requires:

$$\frac{3}{2}\frac{\alpha}{\Delta}\frac{\Pi}{rF}t\,(1 - F(t)) - 2\frac{d}{m} \geq 0$$

Given the previous derivations, it holds that $t\,(1 - F(t)) = \frac{d}{m}\frac{\bar{t}-t}{1-\alpha\frac{\Pi}{rF}\frac{0.5t}{\Delta}}$. Using this form, a sufficient condition for non-negativity of the fourth term is $\bar{t} \geq 2\frac{\Delta}{\alpha}\frac{rF}{\Pi}$, which holds true on $\alpha \in \left[1, \alpha^{thr}\right]$. It could be shown that each of the terms captured in the most outer brackets is non-negative, with at least one term being strictly positive. Thus, the total effect is unambiguously negative.

References

Rosanne Altshuler and Timothy J Goodspeed. Follow the leader? evidence on european and us tax competition. *Public Finance Review*, 43(4):485–504, 2015.

BEPS Monitoring Group. Response to action 5: Harmful tax practices: Agreement on the modified nexus approach, 2015. URL https://static1.squarespace.com/static/5a64c4f39f8dceb7a9159745/t/5af49dbd352f535071a00718/1525980608070/ap5-htps-modified-substance.pdf. last visited: August 29, 2019.

Kurt R Brekke, Armando J Garcia Pires, Dirk Schindler, and Guttorm Schjelderup. Capital taxation and imperfect competition: Ace vs. cbit. *Journal of Public Economics*, 147:1–15, 2017.

Michael P Devereux and Simon Loretz. What do we know about corporate tax competition? *National Tax Journal*, 66(3):745–774, 2013.

Sebastian Eichfelder, Frank Hechtner, and Jochen Hundsdoerfer. Formula apportionment: Factor allocation and tax avoidance. *European Accounting Review*, 27(4):649–681, 2018.

May Elsayyad and Kai A Konrad. Fighting multiple tax havens. *Journal of international Economics*, 86(2): 295–305, 2012.

Roger Gordon and John D Wilson. An examination of multijurisdictional corporate income taxation under formula apportionment. *Econometrica: Journal of the Econometric Society*, pages 1357–1373, 1986.

Andreas Haufler and Marco Runkel. Firms' financial choices and thin capitalization rules under corporate tax competition. *European Economic Review*, 56(6):1087–1103, 2012.

Andreas Haufler, Mohammed Mardan, and Dirk Schindler. Double tax discrimination to attract fdi and fight profit shifting: The role of cfc rules. *Journal of International Economics*, 114:25–43, 2018.

Steffen Juranek, Dirk Schindler, and Guttorm Schjelderup. Transfer pricing regulation and taxation of royalty payments. *Journal of Public Economic Theory*, 20(1):67–84, 2018.

M Karnitschnig and R Van Daalen. Business-friendly bureaucrat helped build tax haven in luxembourg. *Wall Street TO Journal*, 21, 2014.

Michael Keen and Kai A Konrad. The theory of international tax competition and coordination. In *Handbook of public economics*, volume 5, pages 257–328. Elsevier, 2013.

Nicolas Marceau, Steeve Mongrain, and John D Wilson. Why do most countries set high tax rates on capital? *Journal of International Economics*, 80(2):249–259, 2010.

Chakravarthi Narasimhan. Competitive promotional strategies. *Journal of business*, pages 427–449, 1988.

OECD. Action 5: Agreement on modified nexus approach for ip regimes. Technical report, Organisation for Economic Co-operation and Development (OECD), 2015.

Mancur Olson. *The logic of collective action*, volume 124. Harvard University Press, 2009.

Nadine Riedel. Quantifying international tax avoidance: A review of the academic literature. *Review of Economics*, 69(2):169–181, 2018.

Hal R Varian. A model of sales. *The American Economic Review*, 70(4):651–659, 1980.

Part II

Tax Housing or Land?

Tax Housing or Land?
Distributional Effects of Property Taxation in Germany

Rafael Barbosa and Simon Skipka*

Abstract

Despite its theoretical merits, Land Value Taxation (LVT) is not a common policy instrument in most countries. One of the main reasons is uncertainty regarding its distributional impacts. This uncertainty has not been settled by the literature, due to a lack of appropriate data at the household level. We overcome this obstacle by the construction of a unique household level dataset for a sample of German homeowners in 2017. The data collected allows us to study the differences in distributions of land and property values and the resulting distributional effects of implementing a LVT compared to a standard property tax. Our results are as follows. First, we find revenue neutral LVT rates to be around 0.6% in our sample. Second, we find the share of land value in property value on average to be 33% with considerable household heterogeneity, both within and between regions. Third, we find a LVT to be equally progressive if implemented at the federal level, but less progressive if implemented at the regional level, since, although land values are more concentrated than property values, they are not as strongly correlated with income. Quantitatively a revenue-neutral reform from a standard property tax to a LVT at the regional level would increase the average tax burden of the lowest income quintile of homeowners by 25%.

Keywords: Land, Housing, Land Value Taxation, Property Taxation, Distributional Assessment

*Barbosa: Department of Economics, European University Institute, Florence, Italy (email: rafael.barbosa@eui.eu); Skipka: Department of Economics, European University Institute, Florence, Italy (email: simon.skipka@eui.eu). We thank Denis Gorea, Ramon Marimon and Dominik Sachs as well as seminar participants at the EUI for helpful comments and discussions. Further, we thank everybody at the SOEP for their kindness and support during the research visit in Berlin. Finally, we thank Andrea Galeotti for his financial support.

1 Introduction

Land value taxation is a type of property taxation which falls solely on the unimproved value of land, as opposed to standard property tax regimes which take the total value of property (unimproved land plus structures built upon it) as tax base. Since Adam Smith, numerous economists have pointed to the benefits of land value taxation (LVT) over standard property taxation. Most importantly, land, being physically more inelastic than housing, provides a far less distortionary source of revenue for governments.

Despite its theoretical merits, LVT is not widely used. In 2019, from the 36 OECD countries only Denmark, Estonia and Lithuania levy a nationwide LVT. One reason for the small number of LVT regimes is that, historically, a standard property tax has been perceived as more progressive. The idea behind it is that standard property taxation includes an additional part of a household's wealth in the tax base, in the form of structures. Due to this additional component, property values are commonly perceived as a better tag for prosperity than land values. However, in the last years this view has been under scrutiny in academia.

In particular, the issue of land value appreciation has been identified as one of the main drivers of increasing inequality. Rognlie (2015) demonstrates that Piketty's increasing share of capital income is, for the most part, a consequence of increasing value of housing. Furthermore, as pointed out by Stiglitz (2015), this trend in housing wealth is driven primarily by the location premium, rather than by increases in construction costs, meaning land value accounts for an increasing share of total housing value and that its distribution is becoming increasingly more unequal. These findings suggest taxing land values could, in general, be an effective way to redistribute from the top of the wealth distribution to the rest.

Empirical evidence on this matter is scarce and inconclusive. Some attempts have been made to evaluate the distributional impact of a LVT in different US metropolitan areas. The studies find conflicting evidence and suffer under data limitation problems. In this regard, the empirical literature on the distributional impact of a LVT is not yet settled. Our project contributes to this literature, offering fresh evidence from a unique data set for one of the bigger OECD countries, Germany.

Our newly collected data set is superior to previous sources on two dimensions. First, the literature has so far relied on own estimation techniques to determine land values. However, this estimation is ambiguous, especially if the market for unimproved lots is thin. Thus, if a researcher's estimation procedure structurally differs from the official procedure, a distributional assessment will naturally be biased. We overcome this problem by using official land value estimates directly. Whatever estimation procedure is underlying our land values, it will have a direct distributional impact. Second, so far there exists no data set linking households' land values to other characteristics like income. We overcome this problem by using a geo-match approach to add data on land values to a high quality household survey containing a representative sample of German households.

Based on our data set, we provide (new) answers to the following research questions: How do holdings of land and property values differ across the population? How do the differences relate to income? How do the differences vary across and within regions? What is the quantitative impact of these differences when comparing a standard property tax regime to one based on land values?

To address our first research question, we analyze the distribution of land and property value independently. The differences between both distributions are significant. In particular, land values are much more concentrated than property values. While for property values the Gini-coefficient is 0.35, for land values we find the coefficient to be 0.48. These results suggest the distribution of tax burdens would be much less even under a LVT.

In a next step, we compute the land to property value ratio for each household in our sample. We demonstrate that a household's *Land Value Share*, in relation to the average Land Value Share, is a sufficient statistic to determine winners and losers from a revenue-neutral switch to a LVT. A household with a Land Value Share lower than the average wins under a LVT regime and vice versa. Accordingly, if the Land Value Share was the same across households, a switch in tax regimes would not trigger any change in tax burden. We find the distribution of the Land Value Share exhibits a sizable variance around the sample mean of 0.33, showing a switch in the tax regime triggers significant changes in burden for a large part of the population. In numbers, tax burden differs by at least 22% for half of the population.

Relating those differences in tax burden to household income constitutes the distributional assessment of a LVT, the main objective of this paper. Initially, it is important to state that our data clearly shows both LVT and property taxation to be progressive in nature. So, the distributional analysis is designed to determine which of the two is more progressive.

In this analysis, we initially establish a relation between income and the Land Value Share. Income has an effect on the share through its impact on land and structure value. In fact, it can be shown that, if the income elasticity of structure value exceeds the income elasticity of land value, the income elasticity of the Land Value Share is negative. In such cases, qualitatively, a switch to LVT would be regressive.

A prevalent characteristic of property taxation is its regional scope. In most countries property taxes are levied on a sub-federal level, in Germany it is one of the most important municipal taxes. For this reason, our main analysis works under the assumption of *regional* revenue-neutrality. Imposing that restriction we estimate an income elasticity of land value of 0.2 and an income elasticity of structure value of 0.35. Accordingly, the income elasticity of the share is -0.15, indicating a regressive impact of a LVT, compared to a property tax.

An intuition behind this result is that structure value is easier to adjust than land value. Changing the land value by altering the plot size is oftentimes not feasible due to physical constraints, changing the land value by moving to a different neighborhood triggers moving costs. In general, the argument applies independent of the direction of adjustment. However, in particular for the moving we find an accentuated downward rigidity. This means although it is difficult to find high income households in low land value neighbourhoods, it is not uncommon to find low income households in high land value neighbourhoods, living in houses with low structure value.

Quantitatively, our results show that, on average, implementing a LVT rather than a property tax reduces the difference in yearly tax burden between the first and fifth income quintile by 90.€ This number must be seen through the lenses of traditionally low property tax rates in Germany. We find revenue neutral land value tax rates in our data to be 0.6% on average, translating into an average annual tax burden of 800.€ Setting the change in relation to the property tax burden, we find households in the first income quintile experience, on average, a tax burden increase of around 25%. So, in the current context of rising property tax rates, our results have a relevant quantitative impact for future policy debates.

Finally, beyond the average, we find a significant dispersion in tax burdens within each income quintile. For example in the lowest income quintile, despite an average increase of tax burden by 25%, the tax burden under a LVT is smaller for 45% of the households. This heterogeneity shows that income by itself should not be taken as a reliable predictor of the difference in tax burden.

After having laid out our results, we would like to point out that our project arrives at a time in which LVT has become a hot topic in the policy debate.

First, over the last decades several OECD countries, including Germany, have been facing a housing crisis with rents and house prices in major cities soaring. Several reports, such as Mirrlees et al. (2011) and the OECD Fiscal Federalism report of 2016 point to LVT as a promising tool to curb rent hikes by promoting housing construction while also efficiently raising revenue for the government.

Second, the use of German data for this project is well timed as property taxation in Germany is under a process of major reform. In early 2018, the German constitutional court has ruled the property tax must be replaced. Meanwhile, economic research institutes have pointed to a LVT as a possible instrument to supersede it. Additionally, the German government has recently set goals to sharply increase housing supply in the coming years to fight the skyrocketing property prices and rents hurting households in the country. This project can provide a basis to evaluate the suitability of different policy alternatives in achieving these goals. Further, in the latter part of this paper we provide a brief analysis of a LVT's distributional impact when moving away from the current system of property taxation in Germany.

This paper contributes to the literature on several fronts. The main contributions are the empirical identification of the household level distribution of land and property values in a major developed country and the estimation of the distributional impacts of land value taxation and property taxation in relation to their progressivity.

Besides the main contribution, we identify other relevant secondary contributions. First, by building a novel household level dataset with information on property and land holdings. Second, by computing relevant measures of land value at the regional level, enabling the estimation of an taxable land value and revenue neutral land value tax rates in Germany. Finally, through the policy experiments carried out, this project contributes to the ongoing discussion on the reformulation of the property tax system in Germany, by shedding light on the likely winners and losers of different types of property taxation (assuming the current level of revenues is to be maintained).

The rest of the paper is structured as follows. In Section 2 we briefly discuss the literature on distributional aspects of land value taxation. Section 3 explains the construction of the data set used in our analysis. Section 4 presents relevant regional level results. Section 5 contains the distributional assessment at the household level. Section 6 discusses the distributional impact of a LVT moving away from the current property tax regime in Germany. Finally, Section 7 concludes.

2 Literature Review

Theoretical literature addressing the efficiency gains of implementing a LVT is relatively abundant. Property taxation contains an implicit tax on capital which hinders the accumulation of housing capital, creating an inefficiently low level of housing supply in the economy. In contrast, LVT taxes an asset in (quasi) fixed supply, so that a switch in tax base would remove the physical distortion. Aura and Davidoff (2012) for example, show

how optimal property tax rates increase with the share of pure land rents to structures.

Empirically, there have been some attempts to assess the impact on housing supply of switching from a property tax to a LVT. These papers usually rely on using policy changes in specific cities where property taxation follows a two-rate system, taxing land and structures at different rates. Oates and Schwab (1997) focuses on the case of Pittsburgh in the US during the 1980s. Results show strong evidence that switching towards more land value taxation increases construction and overall housing supply against a control group of other cities with similar characteristics, corroborating theoretical arguments.

Few attempts have been made to quantify the distributional aspects of taxing land values instead of property values. England and Zhao (2005) and Plummer (2010) study changes in two-rate property tax systems and find conflicting results regarding the progressivity of the measure. The former finds evidence for a regressive tendency in the case of New Hamphshire, and the latter finds moving to a LVT in Texas would be slightly more progressive, while also shifting the tax burden away from single-family properties and unto other property classes. However, these papers rely on regional level data and thus are unable to pick up on cases of low income households in high income regions or vice-versa.

During the discussion about the introduction of a LVT in Germany, several policy reports stressed the importance of distributional consequences and provided intial evidence. A recent example of this kind is Fuest et al. (2018). The authors discuss distributional consequences between households living in multiple and single family houses, showing a LVT shifts a significant portion of the tax burden to single family house owners. Their study assumes representative type of houses, so that they cannot discuss the idiosyncratic differences in quality and size. Further, the authors are not able to quantitatively link the propensity of living in a certain type of house to a household's income.

3 Data

This paper aims first to fill a gap in the literature by using German data to tease out the distributions of land and property values at a household level for a sample of homeowner households which is broadly representative of the German population. Furthermore, we analyze how the holdings of these assets correlate with income to assess progressivity. Such a breakdown of total housing wealth in land and structures has not been attempted at a national scale. This section lays out in detail the construction of our unique data set, which allows us to perform the distributional analysis in the paper.

3.1 Household Survey

The socioeconomic panel (SOEP) is a German household survey conducted by *Deutsches Institut für Wirtschafts-forschung* (DIW). The SOEP provides the basic information on households in our project. We use SOEP data from 2017 (wave 34). For our analysis, the most important variables in the SOEP are those related to income and real estate property. Monthly income is a standard variable in the SOEP included every year. Information about property is less frequent as it is part of a specific wealth module which is only carried out every five years, at last in 2017. In this module, households provide information regarding their wealth holdings, including the value of their primary residence. The information is only provided by owner-occupiers. As primary residence value is a necessary information for the later analysis, we are forced to restrict the sample to owner-occupiers.

The SOEP does not include a decomposition of property value in land and structures value. We use other sources of data to estimate the land component of the property value. In this decomposition, we employ other information from the SOEP, like the number of dwellings within the household's housing structure.

3.2 VALKIS + M

This section introduces the dataset we need to derive the land component of a household's property value, VALKIS + M. It combines information from three different data sources: the German land registry (*Amtliches Liegenschaftskataster*), the official dataset on land values (*Bodenrichtwerte*), the German statistical offices' regional data base (*Regionaldatenbank des Statistischen Bundesamtes*). We introduce the individual parts in isolation and describe, how they are brought together in order to generate VALKIS + M.

3.2.1 ALKIS

ALKIS is the digitized version of the official German land registry. The smallest geographical unit entered in ALKIS is a *lot*. Our analysis proceeds by using the lot as the unit of observation. For each lot, ALKIS contains information on the type of usage as well as the addresses attributed to the lot. The usages range from residential, industrial and commercial land to forests, rivers and streets. An address is attached to a lot for every independent unit of housing that requires postal correspondence. Historically, a lot describes an economic or contextual unit: a river, a street, a piece of residential land owned by an individual. However, over time, this correspondence has been diluted, so that currently ALKIS contains lots with multiple usages, e.g. lots with farmland and residential land, as well as lots with multiple addresses. In order to later account for those incongruities, we keep the information on the number of addresses and the type of usages for all lots. Finally, ALKIS does not contain information on the size or characteristics of any potential structures on the lot.

In sum, we use ALKIS to generate a dataset with lots as the unit of observation. For each lot, we have precise information on the usage as well as the number of addresses. The geographic extent of our dataset spans the whole surface of five German states: Berlin, Hamburg, Niedersachsen, Nordrhein-Westfalen, Thüringen. Data on the remaining states was not available due to data privacy[1]. The states under consideration have a joint population of about 35 million. The sample of states is representative, consisting of metropolitan as well as rural areas and states from former eastern and western Germany.

3.2.2 Official Land Value Data

Bodenrichtwerte are the results of annual assessments conducted by regional councils of real estate experts (*Gutachteraussschüsse für Immobilienwerte*). They are used as measure of land value throughout our project. In Germany, these land values are used frequently by banks to determine the value of a collateral or in insolvency proceedings to assess the wealth of a defaulting debtor. In the context of the current policy debate on the property tax reform in Germany, Bodenrichtwerte are designated to be used as a main source of information to assess a household's future tax burden. The derivation of the land values is twofold.

[1] Each state has specific policies regarding the availability of this data.

First, the regional councils define land value zones as narrow geographical areas for which the land value does not significantly differ within. The split is based on the experience of the council as well as historic and current information on sales prices of property and land. The area of land value zones depends on the heterogeneity of the neighborhood under consideration, however, it rarely spans an area of more than one square kilometer. Second, the regional councils determine the land values per land value zone. Land values are stated separately for agricultural, commercial and residential land[2]. The zone-specific land values are derived from the collection of land and property sales inside a land value zone within the last years. The preferred source of information is the price of unimproved lots. If not available, land values are derived from the price of improved lots, using hedonic price regressions, or the price of unimproved lots in different land value zones with similar characteristics. Figure 1 shows a map of land values in the municipality of Düsseldorf, where one can see the geographical precision of the land value districts.

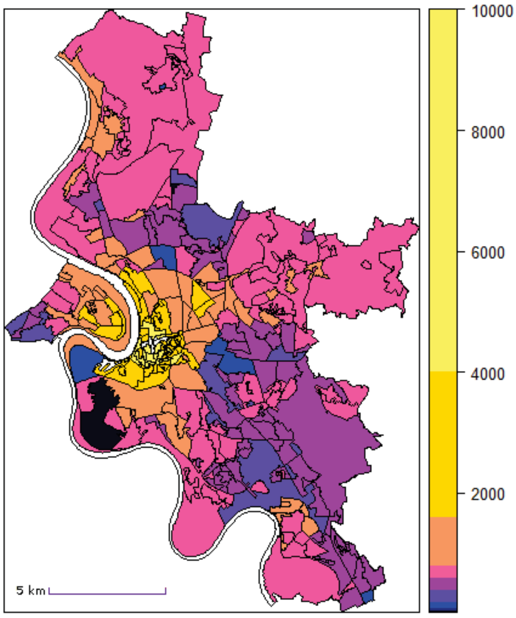


Figure 1: Residential land values in Düsseldorf

Regions in white are non-residential. Values are in Euros per square meter.

In sum, we use the official land value statistics to generate a data set with land value zones as unit of observation. Each land value zone is defined so that land values within a zone do not significantly differ. Further, the data set contains information on different kind of land values within a zone: agricultural, commercial and residential land value.

3.2.3 VALKIS

VALKIS is the result of a spatial joint of ALKIS and the official land value data. The unit of observation is the lot. In particular, we take each lot from ALKIS and find the corresponding land value zone in the official land value data[3]. Conditional on the lot's actual usages, captured in ALKIS, we attach the relevant zonal land

[2]In certain cases, the land values even differ for residential land used for the construction of single or multiple family houses.

[3]The correspondence is given by the spatial reference of both data sets and executed using standard spatial techniques of the statistical software program R.

values to the lot, agricultural, commercial or residential. In sum, VALKIS is a geo-referenced data set with the lot as the unit of observation, the information per lot is: the actual usages, the number of addresses, the land value per m^2 for every type of lot usage.

3.2.4 + M

+ M summarizes information on the regional level, the unit of observation is the municipality. The data is collected from different sources and reflects the living conditions in a municipality in terms of amenities, prices and taxes.

A municipality's degree of urbanization is proxied by population density, a municipality's recent trend in attractiveness, by population growth between 2012-2017. Data on both is gathered from the German statistical offices' regional database. In order to get information on the price level, especially with regard to land prices, we determine the average land value within a municipality. For that, we take the average of a municipality's zonal land values, weighed by the share of residential land contained in a zone[4]. At last, we determine the revenue neutral land value tax rates. Any form of property taxation has traditionally been a municipal tax in Germany and will certainly remain so after the coming reform. Thus, the land value tax rates have to be chosen to guarantee revenue neutrality on the level of the municipality. If we denote by τ_i the revenue neutral tax rate in municipality i, it is defined by $\tau_i \times LV_i = TR_i$, where TR_i represents the current tax revenues and LV_i the aggregate land value of municipality i. Rearranging, the revenue neutral tax rate is given by $\tau_i = \frac{TR_i}{LV_i}$. We can derive the denominator, using the information stored in VALKIS. Regarding the numerator, we once again gather information from the regional database.

Finally, we spatially join VALKIS and +M. The final output is the geo-referenced data set VALKIS + M with the lot as the unit of observation, the information per lot is: the actual usages, the number of addresses, the land value per m^2 for every type of lot usage, Population Density, Population Growth, Average Land Value, Revenue-neutral land value tax rate.

3.3 SOEP 2.0

SOEP 2.0 is the product of a spatial join of SOEP and VALKIS+M, using the SOEPgeo dimension. This unique feature of the SOEP allows us to access the geo-coordinate of each household in the survey. The access is tightly regulated and must be carried out in the DIW facilities in Berlin. We use SOEPgeo to identify the lot in which a household lives and append the respective lot data from VALKIS+M to the original household survey data.

In addition to combining the information, we create additional variables which require the use of data from both of the sources. A crucial variable is the residential size per household. To construct the variable, we take the full lot residential size and divide it by the number of addresses in that lot, from the ALKIS. We further divide by the number of households in each address, which we obtain from the SOEP, to obtain the residential size per household. To exemplify, let's take the total residential size of the lot to be 1000 m^2. Then, if, for

[4]For later sensitivity analysis we generate a second measure of average land value, using the weighed zonal land values within 20km distance of each lot.

example, in the ALKIS the lot is associated with two addresses (two independent residential buildings), and if in the SOEP we observe there are four households in the building, we impute the residential size of our particular household in our sample to be $\frac{1.000m^2}{2\times4} = 125m^2$.

The computation relies on two assumptions, which should be addressed. First, splitting lot land size by the number of addresses in the lot assumes that, in case of multiple addresses, each address occupies an equal fraction of the lot's size. Second, we assume that for multiple family houses, households share the residential size equally. Although these assumptions will lead to errors in specific cases, both reflect the benchmark in the German housing market and, thus, should not influence overall results. The land value component of a household's property value then derives as the product of residential size and residential land value per m^2.

In sum, SOEP 2.0 is a data set with the household as the unit of observation. It carries the variables from the SOEP and augments them with a decomposition of property value in land and structure value. Further, for each household it adds regional information on: Average Land Value, Population Density and Growth, revenue neutral land value tax rate[5].

3.3.1 Quality of Matching

This section discusses the reliability of our geo-match approach in determining a household's land value component. The fact that our final data set was built from several unrelated sources, each with its own shortcomings, and using a self designed geographical matching algorithm, might raise doubts regarding the validity of our SOEP 2.0 data. We try to address such concerns by evaluating if the relation between self reported property values and imputed land values are consistent with each other.

Given that property value is the sum of land and structure value, an increase of one euro in land value, keeping constant the structure value, should imply an increase of one euro in property value. Thus, if our matching is accurate, we should be able to observe this relation in our sample. To test this hypothesis we run a regression of property value (from the SOEP) on the land value we imputed, controlling for structure value. We do not have a variable of structure value in the survey data. If we did, computing the land value component would have been trivial. Instead, we proxy structure value using SOEP variables with information on the quantity and quality of structures: size of the house (in m^2), and condition of the house (a categorical variable with four levels). We run the following model: $PV_i = \beta_0 + \beta_1 LV_i + \beta_2 size_i + \beta_3 condit_i + \epsilon_i$. Results from this regression show a coefficient for β_1 equal to 1.003, not statistically different from one, consistent with our conjecture. This result reassures us regarding the validity of our geo-match approach and the results we will discuss from here onward.

4 Regional Data Analysis

This section provides a summary of the data collected at a regional level, before proceeding to the household level data.

[5] Appendix A contains a diagram representing the construction of SOEP 2.0.

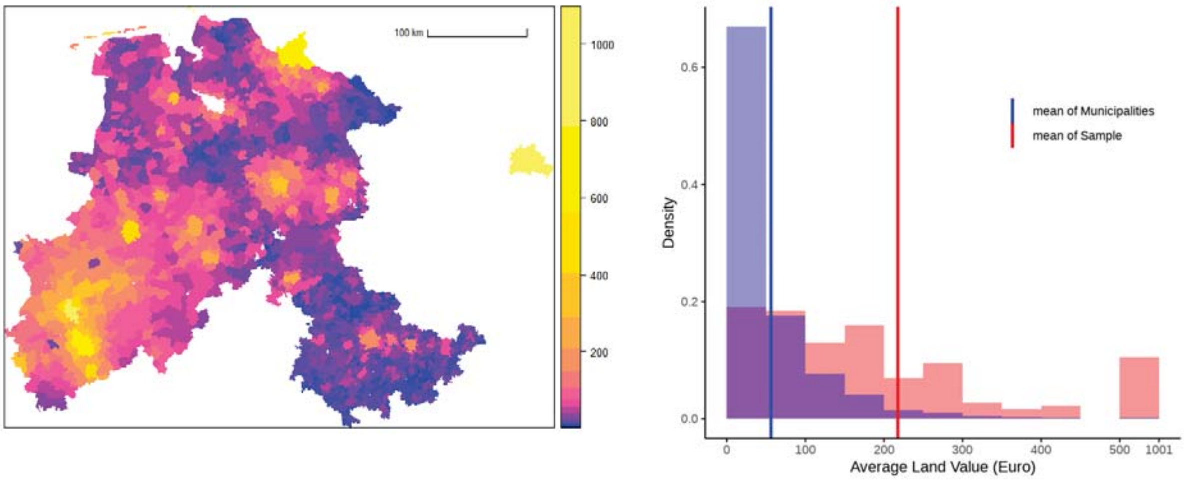


Figure 2: Municipal average land values

The blue (red) distribution in the right panel shows the distribution with municipalities (sample households) as the unit of observation. Values are in Euro per m^2.

The map in Figure 2 shows average land value per municipality (*Gemeinde*) in the five states in our analysis, comprised of a total of 2214 municipalities. It presents a fairly large contiguous region of Germany (apart from Berlin), with different characteristics. The first thing to notice is the heterogeneity in average land values. The lowest municipal average land values in our sample are under 10 €per m^2, while for Berlin (the highest) the average is 1000.€ Very few municipalities exhibit average land values higher than 200,€ as can be seen from the blue distribution in the right panel of figure 2. Nevertheless, a substantial number of observations at the household level are from these municipalities, as can be seen from the red distribution in the same panel.

Our regional data allows for the computation of other interesting aggregate statistics. Total land value in the region we are considering is over 1.5 €trillion, 1.2 times the region's GDP. The magnitude is in line with recent estimates from the US, e.g. Larson (2015). 90% of the total land value is non-agricultural, the rest being agricultural. These numbers establish land value as a sizable, mostly untapped tax base.

Having computed total land values in each municipality and collected the respective current property tax revenues, we have computed the necessary land value tax rates which would ensure revenue neutrality. The histograms of these revenue neutral land value tax rates are presented in Figure 3. Again, in blue the distribution of municipalities, and in red the distribution of households in our sample. Around 70% of municipalities would need to set a tax rate between 0.25 and 1% of land value. The maximum revenue-neutral tax rates we find are around 2%. The household distribution is even more skewed to the left, as a result of more densely populated areas having lower revenue neutral tax rates, on average.

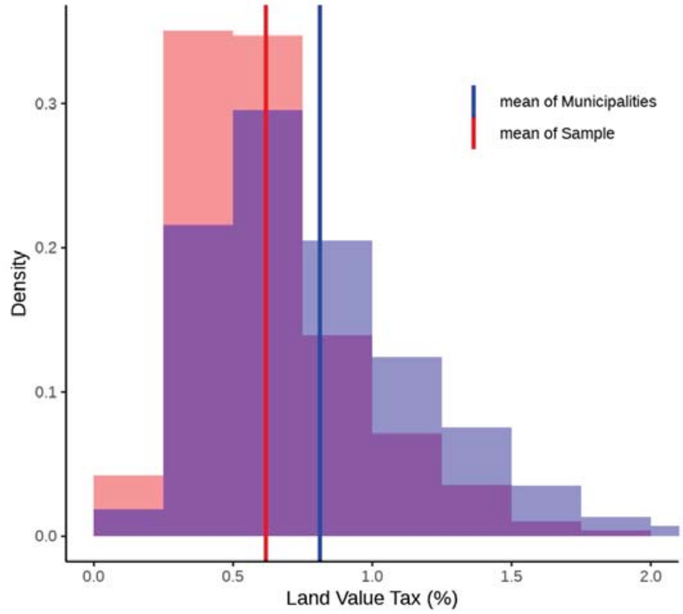


Figure 3: Distribution of revenue neutral land value tax rates

The blue (red) distribution in the right panel shows the distribution with municipalities (sample households) as the unit of observation. The vertical lines represent the mean of each of the distributions.

5 Analysis of the SOEP 2.0

This section contains the main analysis of our paper. We start by presenting the distribution of land and property value in the sample and introduce the concept of *Land Value Share*, which provides a sufficient statistic to qualitatively determine winners and losers from a LVT. We proceed by relating the change in tax burden to income and split up the mechanism in intuitive parts. Finally, the last subsection contains a quantitative assessment of the tax regimes.

5.1 Distributions of Land and Property Value

The first question we address is how the distributions of land and property values differ in our sample. Table 1 provides some initial statistics. Mean property value in our sample is 261.000, € while mean land value is 86.500. € The distribution of land exhibits a higher variance than the distribution of property when controlling for the level of each asset. Standard deviation of property value is 88% the value of the mean, while for land value this number is 124%. Looking at aggregate statistics for total holdings of property and land value in our sample, we see that aggregate land value is 204 €million. This accounts for 33% of aggregate property value, which stands at over 615 €million. The aggregate level of land or property values are important as they represent the size of the tax base of a land value or property tax.

	Property value	Land value	Land value share	Lot size	House size
Mean	260,793	86,495	0.33	603.41	134.14
St. dev	230,018	106,875	0.22	549.76	46.67
Minimum	4,590	980	0.01	7.56	20.00
1st Quartile	150,000	32,640	0.17	255.00	103.00
Median	220,000	58,927	0.27	500.00	126.00
3rd Quartile	300,000	105,300	0.44	779.00	155.00
Maximum	5,000,000	2,536,800	1.19	6,862.00	450.00
Sum	615,210,820	204,042,818			

Table 1: Housing statistics

The sample consists of homeowners in the DIW-SOEP, being residents of the German states Berlin, Hamburg, Lower Saxony, Northrhine-Westfalia und Thuringa. The sample size is 2,359. Lot and House size are in m^2, Land and Property Value are in €.

To assess the concentration of these assets in our sample, we computed Gini coefficients. The value for property is 0.35, while for land it is 0.48. For reference, the Gini coefficient for income is 0.28. It seems land is significantly more concentrated than property in our sample, value wise. If one were to assume the distribution of these assets match the distribution of income on a household level (the household with highest income would also own the most valuable property and land, while the poorest the less valuable property and land), then taxing land would naturally be more progressive than taxing property value. However, this conclusion depends crucially on how these distributions relate to each other and how they relate to income. First, we investigate the link between land and property values.

5.2 Land Value Share

We define the Land Value Share (LVS) as the ratio of land to property value for a given household. This statistic allows one to have a first idea of the magnitude of potential distributional effects. If the distribution were concentrated at a single point there would be no scope for any household to win or lose from a LVT, comparing to a property tax. Dispersion of this measure signifies the existence of households with low (high) land value and high (low) property value, which would thus benefit from paying taxes on their land (property).

In the third column of Table 1 we see the statistics for the LVS. The mean is 0.33 while the standard deviation of this measure is 0.22, a considerably high number, indicating our sample has many households with low property value and high land value, and vice-versa. We can see this more clearly in Figure 4 showing the distribution of the LVS.

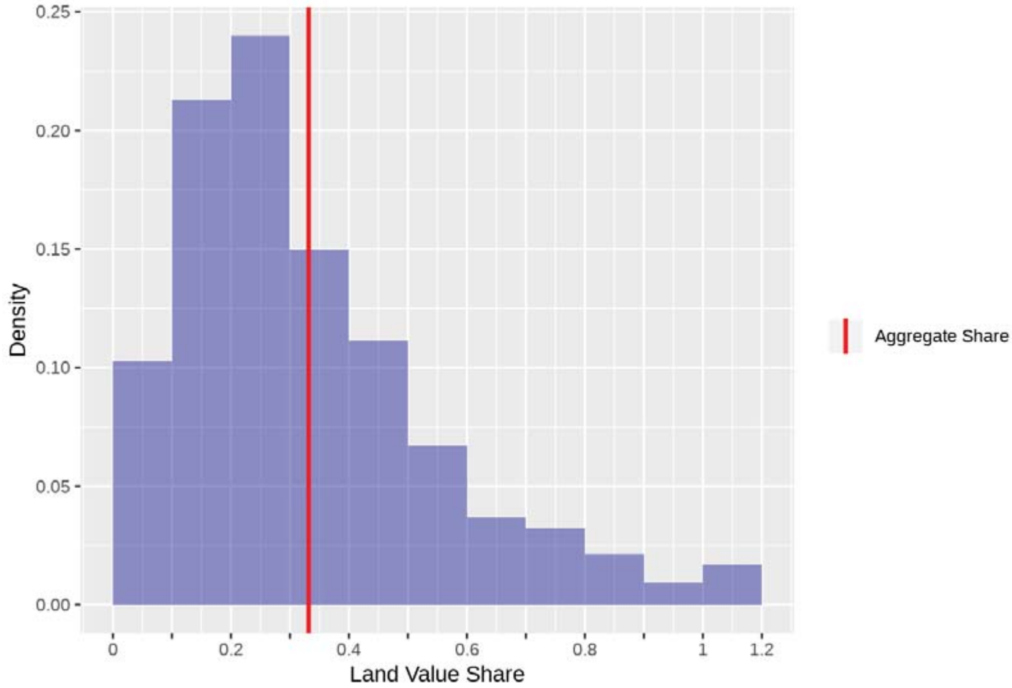


Figure 4: The distribution of land value share

Aggregate share is given by the ratio of total land value to total property value in the sample.

The plot shows the percentage of households in our sample which fall within the bins of LVS we have defined in intervals of 0.1. The distribution is skewed towards lower values, implying the majority of households lives in houses where land value accounts for a relatively low share of property value. Nevertheless, a significant number of households have high LVS as well, even close to 1. The plot shows a mass of around 2% with LVS greater than 1, implying the land is worth more than the property for these observations. While this may appear anomalous, it is entirely possible. A household owning a property which, were it to be sold in the market, would likely imply the demolishing of the existing structures to build new structures, could have a LVS greater than 1 to account for the cost of demolishing.

The vertical line in red depicts the aggregate LVS, meaning the total value of land divided by the total value of property in the sample which can be found in Table 1. This Aggregate LVS (ALVS) will be a centerpiece of the rest of the analysis as it is a crucial threshold defining winners and losers from land value taxation with respect to property value taxation. To understand this, we turn to some simple algebra.

To raise some exogenous level of revenue $\overline{TR}$, the government can choose to either tax land values at a rate τ_L or property values at a rate τ_P, such that $\tau_L \overline{LV} = \overline{TR}$ or $\tau_P \overline{PV} = \overline{TR}$. This means the ratio of the potential tax rates must satisfy

$$\frac{\tau_P}{\tau_L} = \frac{\overline{LV}}{\overline{PV}}$$

At the same time, a household i will pay lower taxes under LVT if $\tau_L LV_i < \tau_P PV_i$. Rearranging and substituting the ratio of tax rates by the ratio of aggregates we get the following condition for a lower tax burden under a LVT

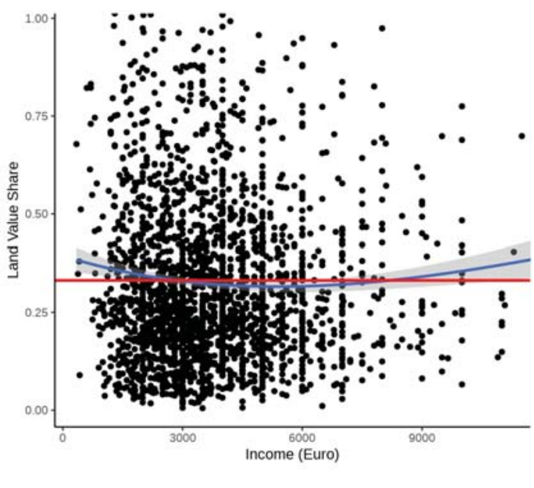
$$\underbrace{\frac{LV_i}{PV_i}}_{\text{LVS}_i} < \frac{\tau_P}{\tau_L} = \underbrace{\frac{\overline{LV}}{\overline{PV}}}_{\text{ALVS}}$$

Households for which $\text{LVS}_i < \text{ALVS}$ (to the left of the red vertical line in Figure 4) will pay less tax under a LVT, those for which $\text{LVS}_i > \text{ALVS}$ (to the right) will pay more. More concretely, this simple result means that if a household owns, for example, a property worth 300.000 €with a land value of 150.000,€ its land value share is 0.5, higher than the ALVS of 0.33. Despite its tax base is only half under a LVT, the household would still pay more, since the levied tax rate has tripled to guarantee revenue-neutrality.

The analysis of the distribution of the LVS reveals that the decision between a LVT or a property tax can create large differences in tax burdens under the different regimes for a substantial number of households. Next, we investigate how our measure of LVS differs with respect to our main characteristic of interest, income.

5.3 Land Value Share and Income

In Figure 5 we see the scatterplot and boxplots of LVS against income and quintiles of income respectively. Also in both plots is the aggregate LVS (in red), separating winners and losers of an LVT. Households below the red line are winners and those above are losers. We see a weak relation between the two. Applying a non-linear trend line reveals the existence of a flat U-shape relation, implying a slight regressive tendency for low income which flips into a slight progressive tendency for higher levels of income.



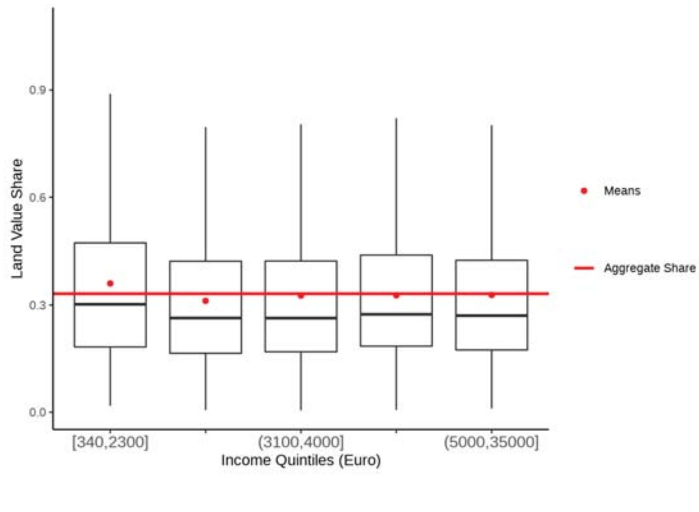


Figure 5: Land value share and income
Income is given as monthly income.

Running a simple OLS regression of the LVS on income proves the weak relation as the coefficient on income is not statistically different from zero. It is important to remember this does not mean LVT is not progressive in itself, only that it is not significantly more or less progressive than a tax on property values. Indeed, a simple regression of land value on income shows a very significantly positive coefficient indicating an increase of 1.000 €in monthly income is associated with an average increase of land value of 14.000 €in our sample.

A weak relation between LVS and income might be surprising, given the previous result showing land values are more concentrated in our sample than property values. An explanation for this would be that, while land is more concentrated, it is less correlated with income than property values. To investigate this hypothesis, we use Figure 6.

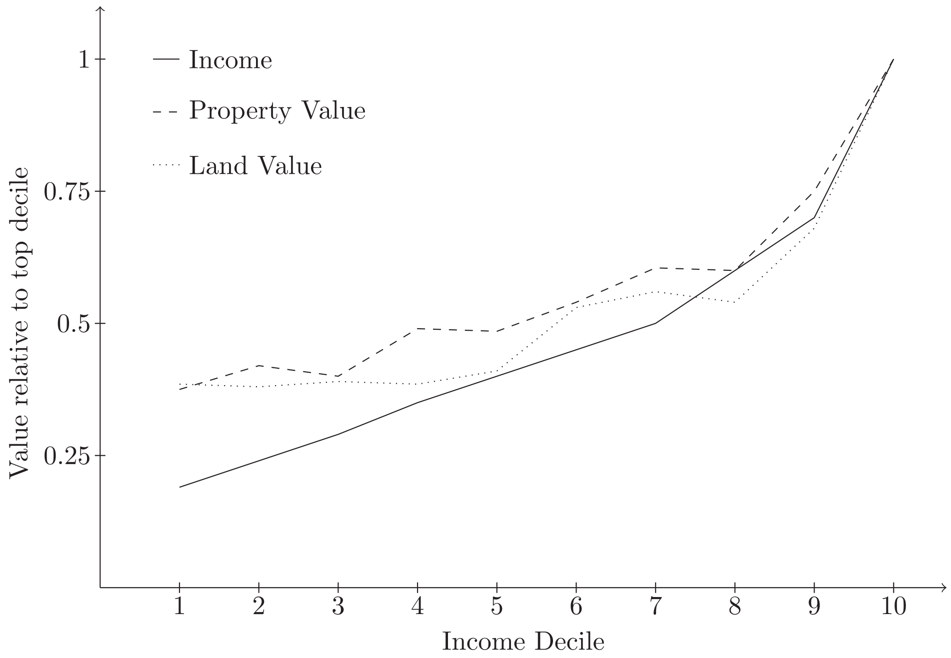


Figure 6: Distribution of income, land value and property value
The graph depicts the decile averages, relative to the value of the 10th decile.

Figure 6 represents how much each income decile holds in average income, land and property value with respect to the holdings of the highest income decile. To exemplify, the plot shows the ninth decile of income on average earns roughly between 60 and 65% of the average earnings in the tenth decile, while holding close to 70% of the value of the property holdings and close to 60% of the value of land holdings. Again, this points to a higher concentration of land values relative to property values. But the most interesting aspect of this plot is how the relative distribution of land values is basically flat for the first five deciles of income. While the fifth income decile earns on average twice as much as the lowest, both have similar levels of land holdings on average. On the other hand, distribution of relative property values exhibits some positive correlation with income even for low income levels. This pattern helps in explaining the flat U-shape found in the relation between LVS and income. For low levels of income, property is a better proxy for income than land, and thus taxing property is slightly more progressive, but in the highest deciles, land is more concentrated than property, so a tax on land values is more progressive as, on average, it hurts top income earners more than a property tax.

5.4 Regional analysis

So far we have been comparing the progressivity of a LVT and a property tax implicitly assuming all households in our sample would be subject to the same rate of each tax no matter where they live, similarly to what would happen if these taxes were levied at a federal level in Germany. However, property taxation is not carried out at a federal, but regional level, more specifically at a municipality (*Gemeinde*) level. For this reason, it is necessary to tailor our analysis accordingly.

Switching from a federal to a regional level analysis poses challenges. Our previous implicit ratio of tax rates was determined by the aggregate land value share in the sample, which is representative on a federal level. Ideally, we would like to do the same at a municipality level, however, for most municipalities we do not have the sufficient number of observations in the sample to reach a meaningful number. As a consequence, working with such a narrow geographical partition is not an option. Instead, we opt to pool municipalities with similar land values by splitting the observations into five quintiles of average municipal land values, in the hope of capturing most of the relevant structural differences. This way, our highest quintile will be comprised mostly of municipalities with the highest average land value (large cities such as Berlin, Hamburg, Düsseldorf, etc.), while the lowest quintile will be comprised of mostly rural municipalities, capturing most of the diverging characteristics of different municipalities. Figure 7 shows a couple of important structural differences across the average land value quintiles. Henceforth, for ease of exposition, these average land value quintiles will be referred simply as land value regions.

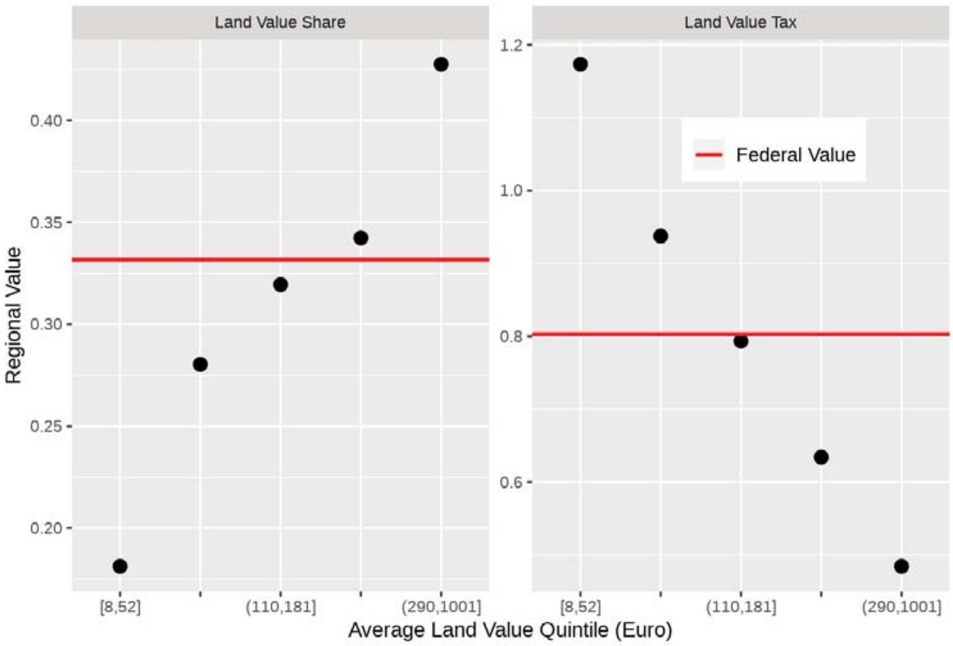


Figure 7: Regional Differences

Panel on the left shows aggregate land value shares computed within each average municipality land value quintile. Panel on the left shows the average revenue neutral land value tax rates within each average municipality land value quintile.

The left panel in Figure 7 shows the aggregate land value share previously discussed at a full sample level (red line), now also computed within each region (black dots). Highest land value region has an aggregate LVS of over 0.45, around 40% higher than the full sample (0.33), and almost three times higher than for the region with lowest average land values (0.16). These differences are decisive for our analysis. A household living in the highest land value region with an individual LVS of 0.4 would be a loser from a LVT implemented at a federal level (as it is above the threshold of 0.33), but would be a winner from a LVT implemented at a regional level (as it is below the relevant threshold of 0.45).

The right panel in Figure 7 shows the heterogeneity of revenue neutral LVT rates across regions. In line with the results in the section on regional differences, regions with higher average land value exhibit lower revenue neutral LVT rates. The highest of the five land value regions has on average a revenue neutral LVT rate below 0.4%, while for the lowest, this number is over 0.8%.

The heterogeneity in regional aggregate land value shares and tax rates indicates there is scope for substantial changes when moving from a federal to a regional analysis. This can be confirmed by a boxplot of LVS across the five land value regions, as seen in Figure 8.

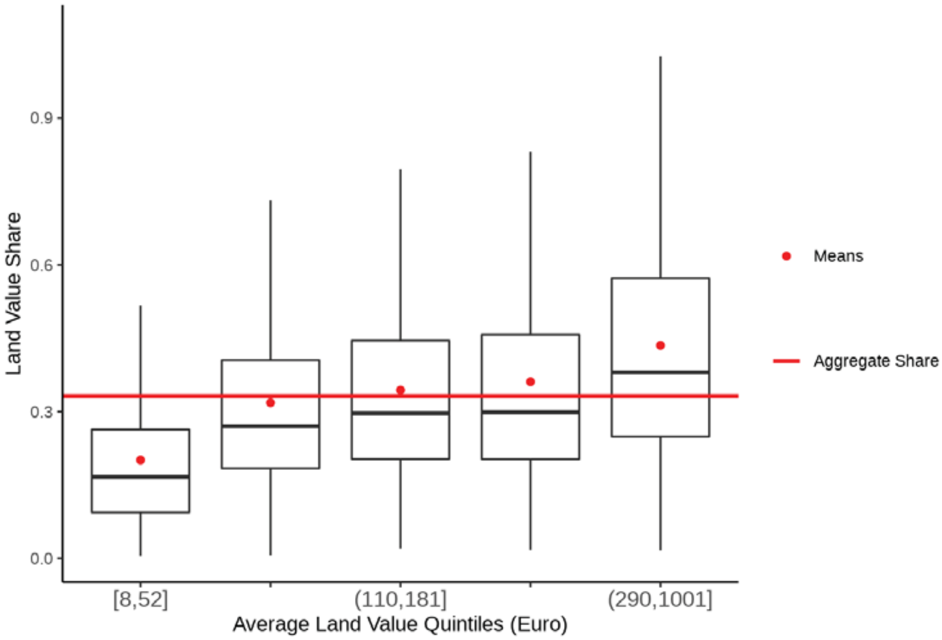


Figure 8: Land value share by average land value

Indeed, Figure 8 reveals stark differences between the two approaches. Using the full sample aggregate LVS (red line) as the threshold to identify winners and losers, one can see a LVT implemented at federal level would lead to more than 75% winners in rural areas (low land value regions) while creating a majority of losers in big cities (high land value regions). Forgetting about the red line and focusing instead on the within region aggregate LVS (red dots), one can see a very different picture, especially for the lowest and highest land value regions. The median of the distribution across the five regions (black line in the boxplot) is below its respective aggregate LVS, indicating more than 50% of households would benefit more from a LVT than a property tax, while with a federal tax the majority of households in cities would lose. Also, the percentage of winners in rural regions is considerably lower, even though more than 50% still win.

Implementing a LVT with a flat rate at a federal level implies substantial inter-regional transfers, from high land value regions to low land value regions. Overall tax neutrality is achieved, but with the burden falling primarily upon big cities. Implementing a LVT at a regional level naturally shuts down the channel of inter-regional transfers as tax neutrality is achieved also at a regional level.

The differences between federal and regional implementation are driven by the strong effect of regional differences in LVS. A log-log OLS regression of LVS on average municipality land value shows a very positive and significant coefficient. An increase of 1% in average land value is associated with an increase of 0.3% of LVS with an R^2 of 16.4%.

At this point, it is natural to ask if the regional implementation of taxes has any impact on the relation between income and the LVS which, with federal taxes, was virtually non-existing. This would imply conducting the analysis while conditioning on the average land value. Table 2 shows the results of a log-log regression of LVS on income including the average land value as a control.

	(1)	(2)
Intercept	-2.781***	-1.599***
	(0.068)	(0.224)
Average Land Value	0.294***	0.310***
	(0.014)	(0.014)
Income		-0.155***
		(0.028)
N	2359	2359
R^2	0.164	0.174
adj. R^2	0.164	1.174

Table 2: Land value share on income and average land value

The table presents the results of log-log OLS regressions. Standard errors in parenthesis.

Results of the regression in Table 2 show that, when controlling for average land value, income has a statistically significant negative impact on LVS. More specifically, an increase of 1% in income is associated with a decrease of 0.15% in land value share in our sample of homeowners. However, it should be noted the inclusion of income in the regression from an initial specification with only average land value modestly increases the R^2, indicating there is a wide dispersion of LVS for households with similar incomes within land value regions and thus that income is not a strong predictor of whether a household will pay more or less under a LVT compared to a property tax.

The change in the coefficient for income after the inclusion of average land value in the regression suggests a positive correlation between income and average land value of the region. In the next section, we will study this relation in greater depth as well as identify all the channels through which income impacts the LVS.

5.5 Decomposing the Income Elasticity of the Land Value Share

In this section we lay out a simple analytic framework to decompose the effect of income on the land value share in several intuitive channels. The decomposition sheds light on the origins of the distributional effect and once again accentuates the importance of a regional consideration. We present our results in terms of elasticities and estimate the main parameters using data from SOEP 2.0.

The LVS of household i is given by $LVS_i = LV_i/PV_i$. Accordingly, income has an impact on the LVS through the denominator (land and structures value) and the nominator (land value). Within the scope of our paper, we keep the effect of income on structure value as a whole, but decompose the effect on land value. Mechanically, we can decompose the household's land value into its constituent components according to our calculation:

$$LV_i = lv_i \frac{lot.size_i}{hh_i}$$

lv_i denotes the land value per m^2, $lot.size_i$ denotes the size of the lot the house of the household is built and hh_i denotes the number of households sharing the lot given by the product of number of addresses and number of neighbours per address. Substituting this expression back into our identity of LVS in logs we get

$$\log(LVS_i) = \log\left(lv_i \frac{lot.size_i}{hh_i}\right) - \log(PV_i)$$

We can further break down this identity until we arrive at a linear relation between the logs of these variables.

$$\log(LVS_i) = \log(lv_i) + \log(size_i) - \log(PV_i)$$

Here, $lot.size_i/hh_i$ was kept as a single variable and renamed $size_i$. In a next step, we break down lv_i into a regional component which is the average land value of the region (Alv_i) and a factor capturing the deviation from the regional average (Rlv_i), which henceforth we denote as relative land value (as in relative to the average of the municipality). So, if household i resides in a lot with a land value per m^2 of $120, €$ located in a municipality where the average land value per m^2 is $100, €$ we can rewrite the 120 as 100×1.2. Applying this decomposition to our LVS expression and once again separating the resulting multiplication inside the log, we arrive at:

$$\log(LVS_i) = \log(Alv_i) + \log(Rlv_i) + \log(size_i) - \log(PV_i) \tag{1}$$

So far, we have decomposed the land value in three components. We continue by setting up a Structural Equation Model (SEM) to quantify the impact of income on LVS through each of them. In order to determine the full impact of each component, we have to quantify their impact through property value, too. We perform the relevant corrections ex post.

From (1) the income elasticity of the share can be decomposed to:

$$\frac{\partial \log(LVS_i)}{\partial \log(I_i)} = \frac{\partial \log(Alv_i)}{\partial \log(I_i)} + \frac{\partial \log(Rlv_i)}{\partial \log(I_i)} + \frac{\partial \log(size_i)}{\partial \log(I_i)} - \frac{\partial \log(PV_i)}{\partial \log(I_i)} \tag{2}$$

The first three terms of (2) are denoted as: Regional Effect (RE), Neighborhood Effect (NE), Size Effect (SE). Broadly, they capture the impact of income on the LVS through: the correlation between the regional price level and income (RE), the decision to live in a neighborhood with a certain level of amenities (NE), the decision to live in a bigger lot and a Single or Multiple Family House (SE). The last term of (2) captures the full impact of income on the LVS through property value and it will be decomposed ex post. Initially, we estimate the individual terms by using the following set of equations in the framework of a SEM:

$$\log(Alv_i) = \alpha_1 + \beta_1 \log(I_i) + \epsilon_{1,i} \tag{3a}$$

$$\log(Rlv_i) = \alpha_2 + \beta_2 \log(I_i) + \gamma_2 \log(Alv_i) + \epsilon_{2,i} \tag{3b}$$

$$\log(size_i) = \alpha_3 + \beta_3 \log(I_i) + \gamma_3 \log(Alv_i) + \epsilon_{3,i} \tag{3c}$$

$$\log(PV_i) = \alpha_4 + \beta_4 \log(I_i) + \gamma_4 \log(Alv_i) + \epsilon_{4,i} \tag{3d}$$

In our SEM-framework, it is important to not only incorporate the direct impact through Alv_i in (3a). In (3b) the inclusion of Alv_i corrects for the fact that in areas with high average land values, mostly cities, the highest land values are measured in zones where residential and commercial usages are mixed, e.g. in city centers. Thus, fewer households live in these zones and so, the relative land value in cities is structurally underestimated. In (3c) the inclusion corrects for the fact that in municipalities with high average land value, mostly cities, the average lot size is structurally smaller. Finally, in (3d) the Alv_i is included to control for different levels of construction costs in cities versus villages.

Using the results of the SEM in (2), the average elasticity is given by:

$$\frac{\partial \log(LVS)}{\partial \log(I)} = (1 + \gamma_2 + \gamma_3 - \gamma_4)\beta_1 + \beta_2 + \beta_3 - \beta_4 \tag{4}$$

The effect through property value, β_4, still carries the effect through land and structure value. We decompose the effect in a structure value effect β_5 and the different land value effects, using the identity $PV_i = SV_i + LV_i$. After some reformulations[6], the structure value effect is given by:

$$\beta_5 = \overline{\left(\frac{PV_i}{SV_i}\right)} (\beta_4 + \gamma_4\beta_1) - \overline{\left(\frac{LV_i}{SV_i}\right)} ((1 + \gamma_2 + \gamma_3)\, \beta_1 + \beta_2 + \beta_3) \tag{5}$$

Using the results in (4), the income elasticity of the LVS finally reads:

$$\frac{\partial \log(LVS)}{\partial \log(I)} = \underbrace{\overline{\left(\frac{SV_i}{PV_i}\right)} (1 + \gamma_2 + \gamma_3)\, \beta_1}_{\text{RE}} + \underbrace{\overline{\left(\frac{SV_i}{PV_i}\right)} \beta_2}_{\text{NE}} + \underbrace{\overline{\left(\frac{SV_i}{PV_i}\right)} \beta_3}_{\text{SE}} - \underbrace{\overline{\left(\frac{SV_i}{PV_i}\right)} \beta_5}_{\text{HE}} \tag{6}$$

The intuition of the first three terms was introduced before. Their magnitude is now corrected for presence in denominator and nominator. The fourth effect is denoted as House Effect (HE). It captures the impact of income on the LVS through the decision to invest in the structure value, by renovation or buildup.

The Regional Effect is a special case in two ways. First, due to simultaneity, the Regional Effect cannot be interpreted causally. Only households with a sufficiently high income can afford to live in cities and surrounding municipalities given the soaring land prices over the last years. However, at the same time, firms in cities tend to pay higher wages in order to compensate for the higher living costs in these areas. Second, as argued in the previous sections, the Regional Effect is irrelevant for a distributional assessment as property taxes are collected on a municipal level.

Our preferred interpretation of the income elasticity of the LVS is the sum of NE, SE and HE, the (regional) net elasticity. However, to accentuate the importance of the regional component and to hinge our analysis to previous sections, we run the full model and present gross and net elasticity separately.

Figure 9 shows the results of our decomposition of the income elasticity of the LVS through a structural equation model. Given the identity-based approach of this section, the estimates of the full elasticities (Gross Elasticity, Net Elasticity) match the results of the log-log OLS regressions of LVS on income previously presented. The gross income elasticity of LVS is not statistically different from zero, while after filtering out the Regional Effect the net income elasticity is -0.15, significantly different from zero.

[6] The reformulation procedure is summarized in Appendix B.

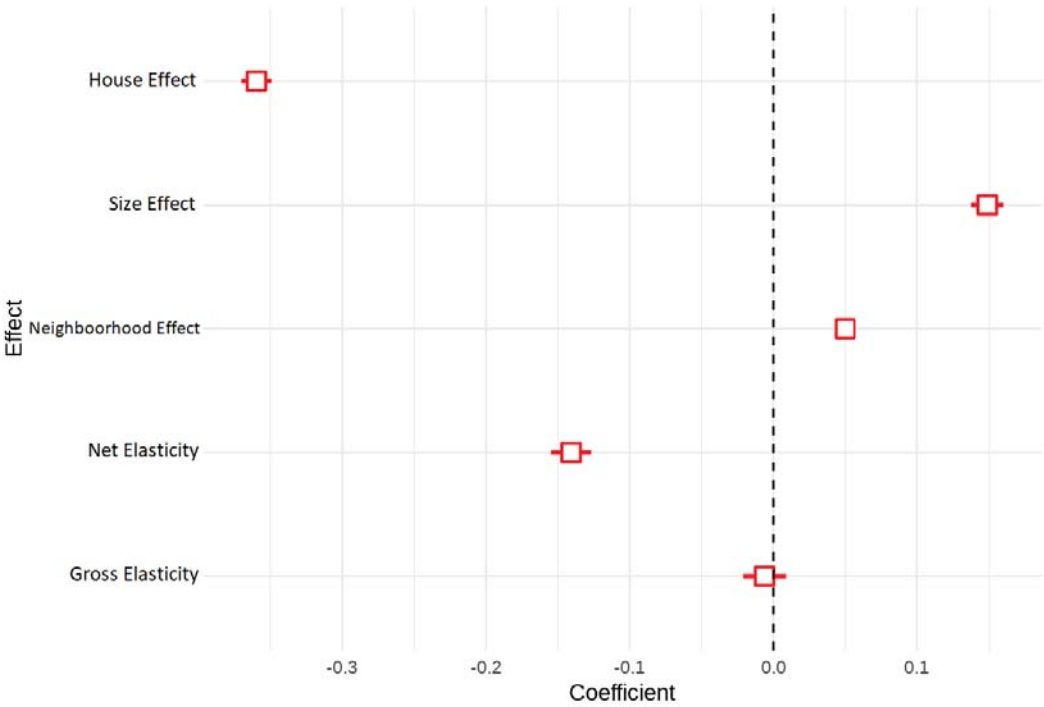


Figure 9: Decomposition of the income elasticity of the land value share.
Gross effect includes effect of income through average land value effect (region effect). Net effect excludes this channel.

We interpret the net effect and its components. The House Effect is close to -0.36, dominating Neighborhood and Size Effect, which are at 0.15 and 0.05. The reason is that structure value is easier to adjust than land value. Changing the land value by altering the lot size is oftentimes not feasible due to physical constraints, changing the land value by moving to a different neighborhood triggers moving costs. In general, the argument applies independent of the direction of adjustment. However, in particular for the Neighborhood Effect we find an accentuated downward rigidity. This means although it is difficult to find high income households in low land value neighbourhoods, it is not uncommon to find low income households in high land value neighbourhoods.

In sum, this section shows that on average households with higher income: occupy larger lots, live in more expensive areas, invest more in renovation and buildup of their houses. Comparing the magnitudes, our analysis reveals that the house margin is the dominant one. Thus property value is a better 'tag' for income than land value, making a Land Value Tax less progressive than a property tax within regions. Finally, to capture this relation it is important to remove the regional veil.

5.6 Quantitative analysis

So far, we have focused our analysis around land value share as a sufficient statistic to determine who wins and loses from a LVT compared to a property tax. However, the LVS hides an important dimension: the magnitude of the change in tax burden. Distance of a household's LVS from the aggregate LVS is not an accurate measure of how much a specific household will be affected. Take two households with the same LVS, which is higher than the aggregate LVS, but where one has values of land and property which are half of the other. Clearly, the household with the highest underlying value of assets stands to lose more from a LVT.

Table 3 summarizes the quantitative impact of a LVT for the different income quintiles.

		Income Quintile				
		1	2	3	4	5
Percentage of losers	in %	54.2	44.8	44.1	42.2	37.2
Mean	in €	39.64	2.80	9.89	-6.64	-49.98
1st Quartile		-78.46	-128.23	-142.79	-159.81	-238.57
Median		21.18	-19.66	-30.45	-46.34	-78.26
3rd Quartile		143.72	102.13	131.10	128.39	109.95

Table 3: Winners and losers of a LVT (I)

The values are computed as the difference between LVT and property tax burden. Positive values indicate higher burden under a LVT.

In general, average LVT burdens range from around 300 €for the lowest income quintile to around 650 €for the highest. Regarding winners and losers, Table 3 picks up the regressive trend we have encountered in previous sections. While over half of the households in the lowest quintile pay more under LVT, this number is 37.2% for the highest income quintile. On the quantitative dimension, the results show that implementing a LVT decreases the difference in the average tax burden between first and fifth income quintile by around 90. €

The quantitative results prove the intuition of our qualitative section, however, the effects turn out to be modest in magnitude. The reason is the traditionally low level of property taxation in Germany. In particular, the revenue neutral land value tax rates have a mean of 0.6%.

The significance of property taxes however has recently risen in Germany. Over the last years, tax rates have increased nationwide. Furthermore, in other countries, property taxation is a much more important source of revenue. Thus, in a next step we provide statistics to show our results potentially will have significant quantitative impact, if the importance of property taxation continues to rise.

In particular, we compute the variation in tax burden as a percentage of the value of one of the tax burdens in order to make it invariant to the scale of the total revenues being raised. This way one can say, for example, household i will pay 30% more under a LVT compared to a property tax. The corresponding monetary burdens depend on the magnitude of the tax rates, but the ratio between the tax burdens would remain unaffected. The results of such analysis are shown in Table 4, again broken into income quintiles.

		Income Quintile				
		1	2	3	4	5
Percentage of losers	in % of Sample	54.2	44.8	44.1	42.2	37.2
Mean	in % of PT Burden	24.49	6.31	8.16	3.88	-4.17
1st Quartile		-34.46	-40.97	-39.76	-40.33	-44.62
Median		8.01	-8.47	-8.31	-12.24	-18.60
3rd Quartile		64.50	36.05	38.51	37.66	22.33

Table 4: Winners and losers of a LVT (II)

The values are computed as the difference between LVT and property tax burden, relative to the property tax burden. Positive values indicate higher burden under LVT.

Table 4 shows considerable differences in tax burdens. The average change in tax burden for the lowest income quintile is 24.49%, meaning households in this quintile would pay, on average, 24.49% more under a LVT than under a property tax regime. For other quintiles, average changes are below 10%. However, the numbers are substantially higher when looking beyond the mean. For more than half of the households in the sample, their burdens change at least 22% under the two different regimes. A quarter of households in the lowest income quintile would pay at least 65% more under a LVT, while another quarter would pay at least 35% less. This analysis confirms our initial assessment that the high dispersion in LVS can lead to significant differences in tax burdens across households.

The data also allows us to investigate in which average land value regions the biggest winners and losers reside. Although one might think the differences would be greater in the highest land value regions, we find the scope and magnitude of the change to be relatively similar across regions.

It is relevant to notice the median voter in our sample of homeowners would be for the implementation of the LVT. A result that holds also within each of the five land value regions we consider. The result that median household pays less under a LVT is a consequence of the higher concentration of land values in our sample of homeowners, leading to a greater share of the total tax burden being paid by fewer households.

6 Current Property Taxation

In Germany the current system of property taxation assesses a household's property tax burden based on property values from 1964 in West- and 1935 in East-Germany. Property values have not been updated subsequently due to the bureaucratic costs of a regular assessment. The outdated property values led the German constitutional court to rule out the current property tax regime in 2018.

In our main analysis we have discussed the distributional consequences of a land value tax as compared to a property tax based on actual property values (NPT). This section complements the analysis by evaluating the distributional consequences of a LVT as compared to the current system of property taxation (CPT). Both type of analyses have attractive features. On the one hand, LVT and NPT are the only realistic options for property taxation in the future. So, a distributional assessment helps to decide on the preferable policy option. On the other hand, for the German case, a comparison of LVT and CPT determines the realized change in tax burden of a LVT, when moving away from the current system of property taxation.

We do not have information on the historic property values used as the tax base in CPT, however, the SOEP contains information on the CPT burden directly. Compared to actual property values, two features of CPT become visible. First, the tax burden increases in the actual property value, as depicted in the left panel of Figure 10. Despite the outdated assessment, CPT captures the differences in property values to a certain extent. Second, the effective average CPT rate decreases in the actual property value, as depicted in the right panel of Figure 10. Compared to the actual property value, CPT is regressive in nature.

Figure 10: Current property tax system

The panels depict the level of current property tax burden (left) and the effective average current property tax rate (right).

Given the historic tax base assessment, CPT is incapable to equalize tax rates on property values, which have diverged since the last assessment in 1964, or 1935 respectively. The main driver for divergence during that period are land values. Differences in the growth rate of land values across municipalities are balanced out by different growth rates in tax rates. However, differences in the growth rate of land values within a municipality necessarily lead to a divergence in effective tax rates. In particular, the effective average tax rate for property values with a relatively high land value is structurally lower under CPT. As most of the high value properties are located in such land value zones, CPT is regressive with respect to actual property values.

Since high value properties are predominantly owned by households with a high income, in isolation, the inability of the current property tax system to equalize effective tax rates on properties across different land value zones within a municipality makes CPT less progressive than a LVT. However, CPT takes into account characteristics of the structures to assess a household's tax burden, which, as argued in our main analysis, makes a LVT less progressive. In total, the mitigation of relative land values in CPT dominates, so that a change from CPT to a LVT would shift tax burden from low to high income quintiles, as depicted in Figure 11.

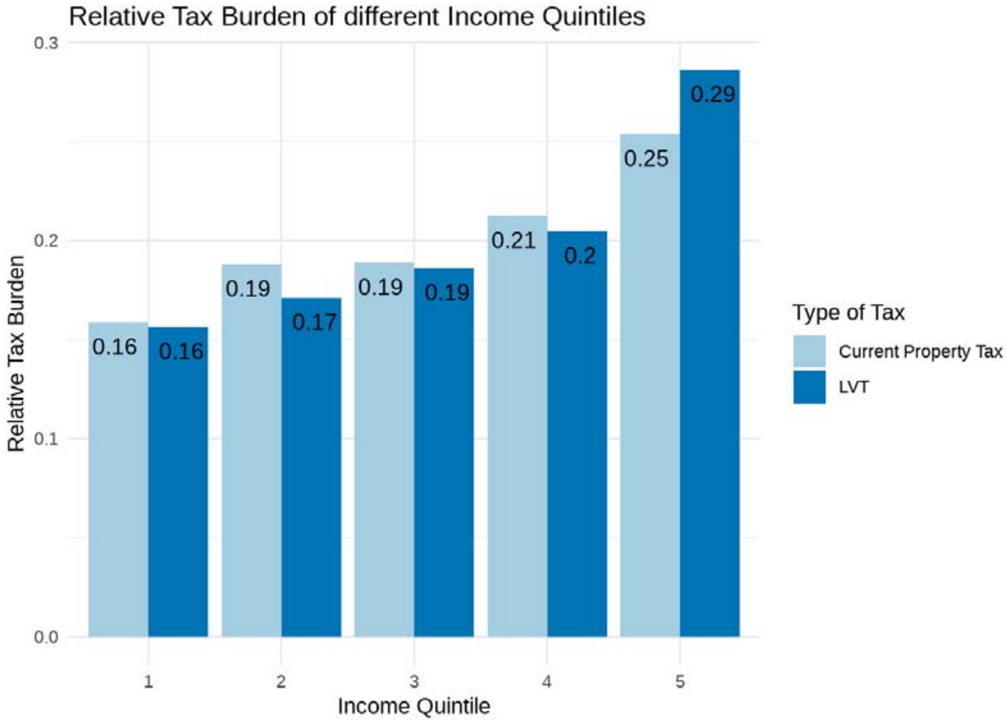


Figure 11: Distributional impact of a LVT compared to the current property tax
The figure depicts the fraction of tax burden carried by the respective income quintiles under the different tax regimes. The land value tax rates are chosen to guarantee budget-neutrality for the sample as a whole.

7 Conclusion

This paper sets out to provide the first empirical identification of the distribution of property and land values at a household level and their relation to income. Those statistics are used to study the distributional consequences of implementing land value taxation compared to standard property taxation. Land value taxation offers various theoretical advantages over property taxation, but the distributional consequences at a household level remained unknown, making its implementation hard to justify. Using geographical matching, official land values and lot data for five German states we successfully estimate the land value associated with the household's primary property value for a sample of close to 2400 homeowners in the German household survey for 2017.

At a municipal level we find revenue neutral property tax rates on average around 0.6%, with considerable regional differences (lower rates the more densely populated). We find the aggregate level of land value to be substantially high, around 1.2 times GDP for the whole region.

At a household level we find considerable heterogeneity in the relative distributions of land and property with an average value of 33% for the share of land value to property value, which was shown to be a sufficient statistic to determine winners and losers from a switch to LVT. We also find no distributional impact from a switch to LVT at a federal implementation level, but a significant regressive impact at a regional level. Given that property taxation has traditionally been a regional tax, the regressive result is our preferred one.

However, the quantitative impact in our sample is modest. Implementing a LVT would decrease the difference of tax burden between the first and fifth income quintile by no more than 100 €annually. This result is due to the traditionally low level of property taxation in Germany. If we set the change in respect to current tax burdens, we find that households of the first income quintile experience an increase of around 25% on average. Thus, in times of rising property tax rates, our results will have an important implication for future policy debates.

Looking ahead, if recent trends of increased gentrification continue, it might lead to an allocation of households across land value areas more in line with household income, making LVT more naturally progressive than property value tax. Regardless, both are likely to produce winners and losers across all income classes, creating the need for careful implementation. This can be accomplished, for example, through exemptions, phase-in periods or the implementation of complementary policies targeted at low income households.

Yet, our results focus solely on primary residences of homeowners. The effect of a switch in the tax regime on renters and landlords critically depends on the incidence of the tax. Standard economic theory suggests a tax on land values would not be passed onto renters, due to the inelastic nature of land supply, making them winners of a LVT. However, in Germany property taxes are traditionally part of the utilities, so that, at least on impact, a part of the change in burden might be passed through to renters. In principle, our data would allow us to incorporate the effects on renters as well. However, such analysis would require additional assumptions on the rent to property value ratio. In order to generate cleaner results, we left renters out of this project. The inclusion of them is an interesting future extension of this project.

This paper focuses solely on the on-impact distributional effects of different types of property taxation, ignoring dynamic general equilibrium effects. If the theoretical benefits of implementing a LVT are realized, namely through higher housing investment, and subsequent control of inflationary pressures on rents, for example, the benefits may outweigh the costs. The construction of a theoretical model which includes these effects is an obvious path for future research on this topic.

A Construction of SOEP 2.0

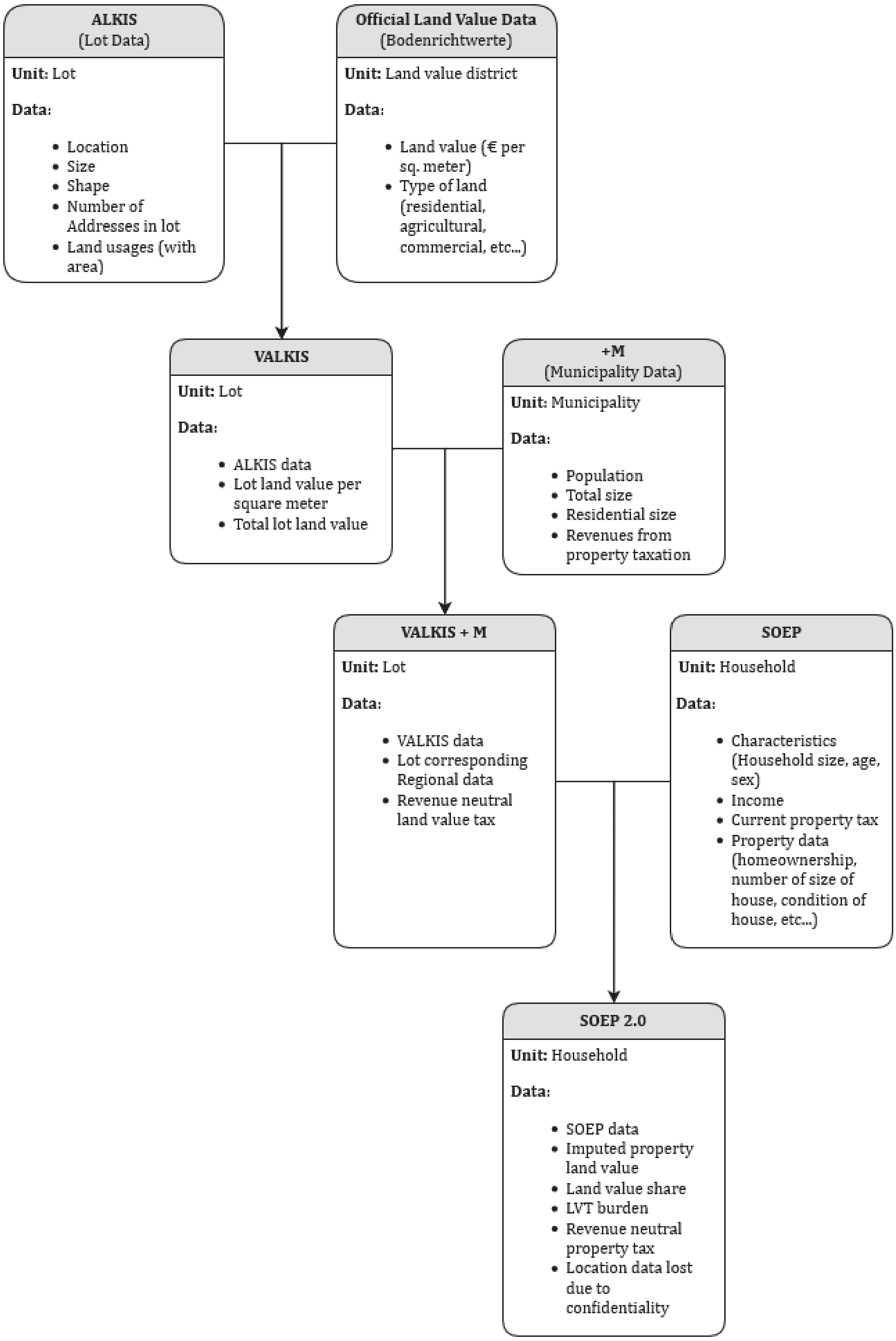


Figure 12: Data flowchart

B Decomposing the Income Elasticity of the Land Value Share

Given the construction of Land Value and Property Value, the following equation holds by identity:

$$\log(\text{LVS}) = \log(\text{Alv}) + \log(\text{Rlv}) + \log(\text{size}) - \log(\text{PV})$$

See that throughout the presentation of the results, we drop the subscripts to ease the exposition.

Accordingly, the income elasticity of the Land Value Share is given by:

$$\frac{\partial \log \mathrm{LVS}}{\partial \log \mathrm{I}} = \frac{\partial \log \mathrm{Alv}}{\partial \log \mathrm{I}} + \frac{\partial \log \mathrm{Rlv}}{\partial \log \mathrm{I}} + \frac{\partial \log \mathrm{size}}{\partial \log \mathrm{I}} - \frac{\partial \log \mathrm{PV}}{\partial \log \mathrm{I}}$$

We can use the results of the regressions (3a) to (3d) in order to reformulate:

$$\frac{\partial \log \mathrm{LVS}}{\partial \log \mathrm{I}} = \beta_1 + \left(\beta_2 + \gamma_2 \frac{\partial \log \mathrm{Alv}}{\partial \log \mathrm{I}}\right) + \left(\beta_3 + \gamma_3 \frac{\partial \log \mathrm{Alv}}{\partial \log \mathrm{I}}\right) - \left(\beta_4 + \gamma_4 \frac{\partial \log \mathrm{Alv}}{\partial \log \mathrm{I}}\right)$$

Once again using the result from regression (3a) that $\frac{\partial \log \mathrm{Alv}}{\partial \log \mathrm{I}} = \beta_1$, we arrive at:

$$\frac{\partial \log \mathrm{LVS}}{\partial \log \mathrm{I}} = (1 + \gamma_2 + \gamma_3 - \gamma_4)\, \beta_1 + \beta_2 + \beta_3 - \beta_4$$

In a next step, we want to decompose the income elasticity of property values in parts, regarding the income elasticity of land value (LV) and structures value (SV). The steps are:

$$\frac{\partial \log \mathrm{PV}}{\partial \log \mathrm{I}} = \frac{\frac{\partial \mathrm{PV}}{\partial \log \mathrm{I}}}{\mathrm{PV}} = \frac{\frac{\partial (\mathrm{SV} + \mathrm{LV})}{\partial \log \mathrm{I}}}{\mathrm{PV}} = \frac{\mathrm{SV}}{\mathrm{PV}} \frac{\frac{\partial \mathrm{SV}}{\partial \log \mathrm{I}}}{\mathrm{SV}} + \frac{\mathrm{LV}}{\mathrm{PV}} \frac{\frac{\partial \mathrm{LV}}{\partial \log \mathrm{I}}}{\mathrm{LV}} = \frac{\mathrm{SV}}{\mathrm{PV}} \frac{\partial \log \mathrm{SV}}{\partial \log \mathrm{I}} + \frac{\mathrm{LV}}{\mathrm{PV}} \frac{\partial \log \mathrm{LV}}{\partial \log \mathrm{I}}$$

Now, define the income elasticity of structures, such that $\beta_5 \equiv \frac{\partial \log \mathrm{SV}}{\partial \log \mathrm{I}}$. Furthermore, by our identities it holds that $\frac{\partial \log \mathrm{LV}}{\partial \log \mathrm{I}} = \frac{\partial \log \mathrm{Alv}}{\partial \log \mathrm{I}} + \frac{\partial \log \mathrm{Rlv}}{\partial \log \mathrm{I}} + \frac{\partial \log \mathrm{size}}{\partial \log \mathrm{I}}$. Using the definitions and the results from (3a) - (3c), we can reformulate:

$$\frac{\partial \log \mathrm{PV}}{\partial \log \mathrm{I}} = \frac{\mathrm{SV}}{\mathrm{PV}} \beta_5 + \frac{\mathrm{LV}}{\mathrm{PV}} \left((1 + \gamma_2 + \gamma_3)\, \beta_1 + \beta_2 + \beta_3\right)$$

Finally, from (3d) we also know that it holds that $\frac{\partial \log \mathrm{PV}}{\partial \log \mathrm{I}} = \beta_4 + \gamma_4 \beta_1$. Putting the equations together, we derive:

$$\beta_4 = \frac{\mathrm{SV}}{\mathrm{PV}} \beta_5 + \frac{\mathrm{LV}}{\mathrm{PV}} \left((1 + \gamma_2 + \gamma_3)\, \beta_1 + \beta_2 + \beta_3\right) - \gamma_4 \beta_1$$

Using this result, the income elasticity of the land value share derives as:

$$\frac{\partial \log \mathrm{LVS}}{\partial \log \mathrm{I}} = \frac{\mathrm{SV}}{\mathrm{PV}} (1 + \gamma_2 + \gamma_3)\, \beta_1 + \frac{\mathrm{SV}}{\mathrm{PV}} \beta_2 + \frac{\mathrm{SV}}{\mathrm{PV}} \beta_3 - \frac{\mathrm{SV}}{\mathrm{PV}} \beta_5$$

In this formulation, β_5 can be recovered from results of (3a) - (3d) and multiplied by $\frac{\mathrm{SV}}{\mathrm{PV}}$ it constitutes the income elasticity of the land value share through the elasticity of the structures value, our house effect.

References

Saku Aura and Thomas Davidoff. An Analysis of Constrained Property Taxes in a Simple Optimal Tax Model. *CESifo Economic Studies*, 58(3):525–543, September 2012. URL https://ideas.repec.org/a/oup/cesifo/v58y2012i3p525-543.html.

Richard W. England and Min Qiang Zhao. Assessing the Distributive Impact of a Revenue–Neutral Shift From a Uniform Property Tax to a Two-Rate Property Tax With a Uniform Credit. *National Tax Journal*, 58(2): 247–260, June 2005. URL https://ideas.repec.org/a/ntj/journl/v58y2005i2p247-60.html.

Clemens Fuest, Lea Immel, Volker Meier, and Florian Neumeier. Reformoptionen für die Grundsteuer B: Eine empirische Analyse der Belastungswirkungen. *ifo Schnelldienst*, 71(22):23–29, November 2018. URL https://ideas.repec.org/a/ces/ifosdt/v71y2018i22p23-29.html.

William Larson. New Estimates of Value of Land of the United States. BEA Working Papers 0120, Bureau of Economic Analysis, April 2015. URL https://ideas.repec.org/p/bea/wpaper/0120.html.

James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba. The mirrlees review: Conclusions and recommendations for reform. *Fiscal Studies*, 32(3):331–359, 2011. URL https://EconPapers.repec.org/RePEc:ifs:fistud:v:32:y:2011:i:3:p:331-359.

Wallace Oates and Robert M. Schwab. The impact of urban land taxation: The pittsburgh experience. *National Tax Journal*, 50(1):1–21, 1997. URL http://EconPapers.repec.org/RePEc:ntj:journl:v:50:y:1997:i:1:p:1-21.

Elizabeth Plummer. Evidence on the distributional effects of a land value tax on residential households. *National Tax Journal*, 63(1):63–92, 2010. URL http://EconPapers.repec.org/RePEc:ntj:journl:v:63:y:2010:i:1:p:63-92.

Matthew Rognlie. Deciphering the fall and rise in the net capital share. *Brookings Papers on Economic Activity*, 46(1 (Spring)):1–69, 2015. URL http://EconPapers.repec.org/RePEc:bin:bpeajo:v:46:y:2015:i:2015-01:p:1-69.

Joseph E. Stiglitz. New Theoretical Perspectives on the Distribution of Income and Wealth among Individuals: Part II: Equilibrium Wealth Distributions. NBER Working Papers 21190, National Bureau of Economic Research, Inc, May 2015. URL https://ideas.repec.org/p/nbr/nberwo/21190.html.

Part III

Mind the Twin Tax Difference

Mind the Twin Tax Difference - Local Tax Competition as an Anti-Coordination Game

Simon Skipka**

Abstract

The literature on local business taxation has so far been focused either on differences in tax rates between neighboring municipalities of different size, the *core-periphery difference*, or on differences in tax rates between municipalities of similar size in different regions, the *regional difference*. Instead, this paper investigates differences in tax rates between neighboring municipalities of similar size, the Twin Tax Difference. First, the paper shows this pattern to be a significant phenomenon in local business taxation: About 35% of the 11,000 German municipalities have a Twin Tax Difference of at least two percentage points. Second, the paper provides a novel theory for the emergence of Twin Tax Differences as an outcome of tax competition between adjacent, fully homogeneous municipalities. If neighboring municipalities share a common labor market, a low neighbor tax rate will exert a negative externality on the ability of a municipality to attract new firms to the region. Concretely, if a municipality attracts firms from outside the region by setting a low tax rate, wages on the regional labor market increase and, thus, a municipality in the neighborhood had to decrease its tax rate even more to attract additional firms. Based on this mechanism, Twin Tax Differences emerge as the equilibrium of an anti-coordination game. Finally, in contrast to the classic negative tax base-externality, the impact of the labor market-externality on welfare is ambiguous in sign, because it increases private consumption in high tax municipalities. The last part of the paper provides empirical evidence on the main assumptions of the model, which are necessary for Twin Tax Differences to emerge.

Keywords: Tax Competition, Local Business Taxation.

**Department of Economics, European University Institute, Florence, Italy (email: simon.skipka@eui.eu).

1 Introduction

Local business taxation has traditionally been a common testing ground for the theoretical predictions of tax competition models. Given that tax laws and other hardly comprehensible regulations are relatively harmonized, corporate tax competition between municipalities or states boils down to the choice of the corporate tax rate, a widespread setting in the tax competition literature[1]. This project sets out to draw attention to a significant empirical pattern on local business tax rates, which has so far been overlooked by the literature: the emergence of *Twin Tax Differences*.

In particular, I define Twin Tax Differences as the differences in tax rates between municipalities of the same size from the same narrow geographic region. First, the quality of a municipality's public goods is mainly determined by its location, e.g. by the accessibility of input markets, or by its size, e.g. by agglomeration effects (Brülhart et al., 2012). Second, the preferences for taxation develop continuously over municipality borders, as it can be seen by highly symmetric electoral results of neighboring municipalities[2]. Accordingly, I consider municipalities from the same region and of the same size to be homogeneous in all dimensions relevant for the choice of the local business tax rate, that is they are *twins* from a public finance point of view.

Looking at data on local business tax rates in German municipalities, we can see how Twin Tax Differences are a relevant phenomenon. By Constitution, German municipalities are guaranteed fiscal autonomy, in particular the freedom to determine their local business tax rates individually. The regulation leaves us with about 11,000 autonomously determined local business tax rates. The data shows that for around 35% of the municipalities there exists a municipality of the same size in the same region with a tax rate different by at least two percentage points. Furthermore, for 10% of the municipalities, there exists a neighbor of the same size with a Twin Tax Difference of more than three percentage points.

Despite its empirical significance, theoretical arguments for the emergence of Twin Tax Differences are scarce. A class of theoretical models able to generate Twin Tax Differences builds on the seminal paper Tiebout (1956). In this class of models, tax competition takes place within a set of ex ante homogeneous municipalities. Prior to the choice of the optimal tax rate, each municipality is populated by a mass of mobile households, who differ in their preferences for public good provision. In equilibrium, municipalities charge different tax rates (Twin Tax Differences emerge) and provide different levels of public goods. The equilibrium can be sustained because mobile households sort across municipalities with respect to their preferences for public goods, and so municipalities differ in their optimal tax rate ex post by a median voter argument. The key of Tiebout models is the sorting of households with respect to their preferences for public goods, which makes municipalities heterogeneous ex post. In the data, sorting would imply heterogeneous electoral results within narrow geographical areas. Given that such patterns are rarely observed in the data, this paper proposes a theoretical model to rationalize Twin Tax Differences, neglecting differences in the preferences for public good provision.

Another reason for the emergence of Twin Tax Differences could be that this paper's definition of twinship, based on location and size, misses other idiosyncratic factors behind the choice of a municipality's tax rate. As an example, recent literature has shown a negative correlation between the fine granulation of a municipality's

[1] For a recent summary of the literature, see Agrawal et al. (2015).

[2] For example, see the vote shares by municipality in the last federal elections in Germany of September 2017: https://interaktiv.morgenpost.de/gemeindekarte-bundestagswahl-2017/.

tax base and the tax rate (Bohm et al., 2016). While not neglecting that a share of the detected Twin Tax Differences can be rationalized by similar argument, given its significance, this paper sets out to provide a structural mechanism to explain the emergence of Twin Tax Differences as an endogenous outcome of tax competition, without the help of any ex ante-heterogeneity.

To rationalize the emergence of Twin Tax Differences, I set up a model of tax competition between homogeneous municipalities. The economic environment incorporates two features, which are particularly important to model competition between adjacent and integrated municipalities. First, some input markets of adjacent municipalities are interrelated. Most importantly, municipal labor markets are integrated, as it can be seen by the high numbers of cross-border commuters. Second, municipalities compete for firms within *and* outside their region. In the case of Germany, Janeba and Osterloh (2013) has documented the presence of extra-regional tax competition by municipalities, conducting a survey among municipal politicians. The theoretical part of this paper picks up these two features and shows the ability of their interaction to explain the emergence of Twin Tax Differences.

In fact, the interaction of integrated municipal labor markets and extra-regional tax competition generates an externality of a municipality's tax rate on wages in adjacent municipalities. Due to the integration of municipal labor markets, wages are determined by the intersection of regional labor demand and regional labor supply. Accordingly, a decrease in the tax rate by a municipality has a positive effect on the regional labor demand if and only if the municipality attracts firms from outside the region. The extent to which the increase in labor demand translates into higher wages on the regional labor market is determined by the labor supply elasticity, which, in its turn, crucially depends on the regional labor market's degree of segregation. If the regional labor market is partly segregated, e.g. due to a limited in-commuting of workers, the region's labor supply is relatively inelastic. In such case, an increase in labor demand in one municipality will increase wages in all municipalities of the region Monte et al. (2018).

Directly, the externality increases welfare in adjacent municipalities by the positive impact on private consumption. At the same time, it diminishes the ability of adjacent municipalities to attract firms from outside the region: The higher the wages on the regional labor market, the lower the tax rate of a municipality has to be in order to attract a certain firm from outside the region.

If a municipality decreases its tax rate, the incentives for an adjacent municipality to adjust its tax rate are two-fold. On the one hand, the municipality has an incentive to decrease its tax rate to limit the outflow of regional firms to the adjacent municipality. On the other hand, the municipality's general incentive to decrease the tax rate diminishes as the ability to attract firms from outside the region deteriorates due to the higher regional wages. If the second effect is strong enough, Twin Tax Differences emerge as the equilibrium outcome of a tax competition game between adjacent, homogeneous municipalities.

An earlier model that discusses municipal tax competition under commuting is Braid (1996). Similar to this paper, the author considers tax competition between adjacent, homogeneous municipalities in an environment in which workers commute at zero costs within a region. The scope of his paper is to discuss optimality of different kind of taxes, e.g. capital or labor tax, in a model with commuting. To keep tractability of his general framework, he restricts the analysis to symmetric equilibria only. My paper proposes a more stylized version of the model, which however allows us to analyze asymmetric equilibria.

Theoretically, Twin Tax Differences were shown to emerge in environments in which regional labor markets are segregated and municipalities have the possibility to attract firms from outside the region. The last part of the paper provides some initial empirical evidence on both characteristics.

An ongoing policy debate is on the impact of decentralized municipalitan taxing rights on regional welfare. In particular, focusing on the implications of decentralization on horizontal tax competition, local business tax competition is generally assumed to generate inefficiently low tax rates Keen and Kotsogiannis (2002)[3]. This project identifies a second horizontal externality, working through the interaction of municipal labor markets, that instead suggests decentralized tax rates to be inefficiently high. In particular, for municipalities with given levels of labor market segregation and intake of extra-regional firms, the second externality potentially dominates and equilibrium tax rates might be indeed too high from a regional viewpoint. In such environment, the provision of targeted grants to foster competition between municipalities might be the optimal policy to increase regional welfare.

The rest of the paper is structured as follows. In Section 2 I present evidence on the significance of Twin Tax Differences, using data on local business taxation in Germany. Section 3 lays out the theoretical model to rationalize the emergence of Twin Tax Differences and Section 4 provides the respective equilibrium analysis. In Section 5 I discuss the main results of the model and its implications on regional welfare and the resulting policy implications. Section 6 presents some empirical evidence on the main assumptions of the model. Section 7 concludes.

2 Stylized Facts

This section presents descriptive statistics on local business tax rates in Germany, to underline the significance of Twin Tax Differences. Local business taxation is the most important source of tax revenues for German municipalities, summing up nationwide to €48 billion in 2017. By Constitution German municipalities are guaranteed fiscal autonomy in the choice of their local business tax rate. Accordingly, Germany provides us with a set of 11,387 local business rates, individually chosen by local policymakers. Due to availability, this section uses data on local business tax rates from 2015.

Two empirical patterns, which have been well documented by previous literature, are differences in tax rates between regions, *regional differences*, and differences in tax rates between municipalities of different size, *size differences*. Regional and size differences can be found in the data on German local business tax rate, as documented in Figure 1.

[3]In their paper, however, the overall implication of decentralized taxing rights is ambiguous due to the existence of a vertical tax externality. This project will abstract from any vertical externality.

Population	Number of Municipalities	Germany	Northeast	NRW	South
		average local business tax rate			
- 100	203	12.9	-	-	12.3
- 200	472	12.4	10.9	-	12.0
- 500	1527	12.4	11.3	-	12.0
- 1000	1816	12.2	11.4	-	11.8
- 2000	1883	12.2	11.5	-	11.8
- 3000	1044	12.3	11.4	-	11.9
- 5000	1186	12.3	11.5	15.9	11.9
- 10,000	1334	12.5	11.8	15.0	11.9
- 20,000	885	13.0	11.8	15.2	12.1
- 50,000	504	13.7	12.8	15.4	12.6
- 100,000	108	14.9	14.7	16.0	13.3
- 200,000	40	15.4	15.7	16.7	14.3
- 500,000	25	16.2	15.9	17.1	15
$>$ 500,000	14	16.0	-	16.5	15.8
Total	11041	12.5	11.5	15.5	11.9

Figure 1: Statistics on Regional and Size Differences

Average local business tax rates are in percentage points. The Northeast consists of the states of Mecklenburg-Vorpommern and Brandenburg. The South consists of the states of Bayern and Baden-Württemberg.

To determine the presence of size differences, the set of municipalities is partitioned in accordance with the official German system of municipality size classes (*Gemeindengrößenklassen*). As we can see, average tax rates are stable across municipality size classes for smaller municipalities. However, once the population exceeds 10,000, the average tax rate increases in community size. There are two explanations for this observation. On the one hand, the per-capita public good expenditure increases in municipality size, since larger municipalities need to spend more resources because of additional administrative tasks or the mitigation of certain congestion problems. On the other hand, larger municipalities are better able to extract location premia from firms, e.g. due to agglomeration effects or positive spillovers from universities.

To sketch the presence of regional differences, the right part of the table is focused on three particular regions of Germany. The northeast region captures the economically weakest parts of Germany. The south, mainland of the German conservative parties, is an economically strong region with preferences for low taxation. Northrhine-Westfalia (NRW), mainland of the German socialdemocratic party, is a region with a preference for high taxation. We can see, in fact, that local business tax rates in NRW are significantly higher than both the Northeast and the Southern part of Germany, which, instead, show comparable average local business tax rates. However, it is important to remark that local business tax rates in the South are low due to the preferences against taxation, whereas in the northeast, low tax rates are chosen to attract investment despite the lower quality of local public goods.

Looking beyond these phenomena, however, there is a significant number of differences in tax rates between municipalities of the same size in the same region. The partition of municipalities by size is again taken from the official approach of the German municipality size classes. Municipalities are considered adjacent if their geographic centres are within a distance of 15 km, which is the average commuting distance in Germany. Based on these rather demanding definitions of location- and size-symmetry, the presence of Twin Tax Differences is documented in Figure 2.

Population	Number of Municipalities	N o M with Neighbors	TTD 0-1	TTD 1-2	TTD 2-3	TTD >3
			in % of municipalities with neighbors			
- 100	203	197	15.2	30.5	20.3	34.0
- 200	472	458	24.2	45.4	23.8	6.6
- 500	1527	1498	10.4	47.3	28.0	14.3
- 1000	1816	1788	18.4	45.6	24.6	11.5
- 2000	1883	1872	20.4	38.6	30.6	10.4
- 3000	1044	1023	33.6	42.4	18.3	5.7
- 5000	1186	1175	22.6	47.9	25.4	4.1
- 10,000	1334	1325	17.2	47.2	26.6	8.9
- 20,000	885	868	16.2	47.7	24.5	11.5
- 50,000	504	466	19.3	46.6	25.5	8.6
- 100,000	108	72	41.7	41.7	8.3	8.3
- 200,000	40	19	63.2	36.8	0	0
- 500,000	25	9	44.4	22.2	33.3	0
> 500,000	14	4	25.0	75.0	0	0
Total	11041	10774	19.7	44.6	25.6	10.1

Figure 2: Statistics on Twin Tax Differences

The difference between the number of municipalities in column 1 and 2 are because of municipalities which do not have an other municipality of the same size within a 15km distance. TTD are computed as follows: Of those municipalities which have a neighbor of the same size, how many have a difference in tax rates of x to y percentage points (in case of multiple neighbors, the highest difference in tax rates was used).

Twin Tax Differences of at least two percentage points or more are present for 35 % of the municipalities and the phenomenon is present across all size classes. For 10 % of the municipalities there even exists a municipality of the same size in the same region with a tax rate different by three percentage points or more. In order to see the economic significance, consider a firm who generates an annual profit of 100,000€ and works under a real return rate of 5%. For such a firm, the discounted stream of excess tax burden from locating in a municipality with a Twin Tax Difference of three percentage points is $\frac{3{,}000€}{0.05} = 60,000€$. Given the existence of moving costs, surely not every firm will find it optimal to move in response to the difference in tax rates. However, especially for those about to change offices anyway, like firms increasing their capacity, or for firms without much installed equipment, like firms from the service sector, such differences could trigger a move to low tax municipalities[4].

After having presented evidence on the significance of Twin Tax Differences, it is relevant to investigate their geographical extent as well, in order to address the question of whether TTD only emerge in particular

[4]For evidence of heterogeneous tax-induced relocation between firms of different age and sector, see Giroud and Rauh (2019).

regions or spread out over the whole of Germany. The evidence on the regional extent of Twin Tax Differences is presented in Figure 3.

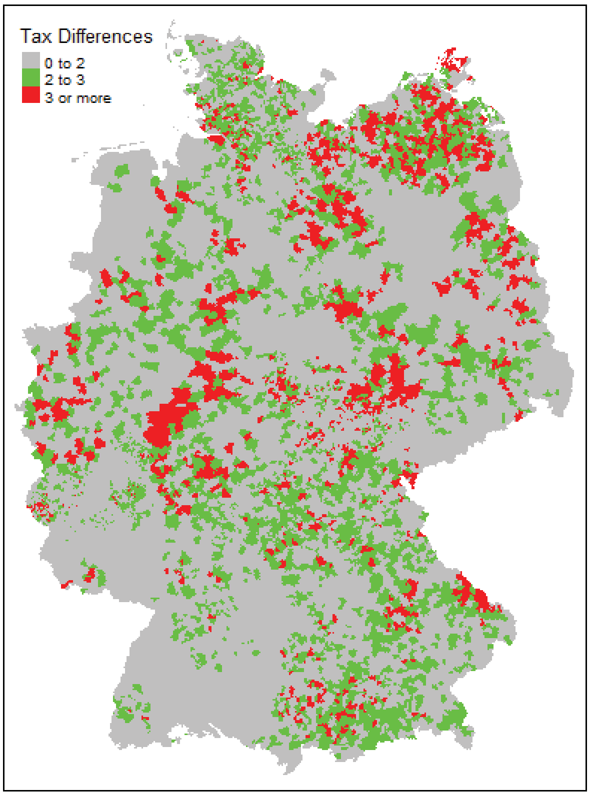


Figure 3: Twin Tax Differences in Germany
Twin Tax Differences are depicted in percentage points.

Twin Tax Differences of more than three percentage points are generally present in many different parts of Germany. However, Twin Tax Differences do not evenly spread out across the country, but are locally concentrated in certain regions, e.g. in the Northeast or the Southeast. In this regard, the theoretical part of this project sets out to identify economic characteristics of a region necessary for the emergence of Twin Tax Differences.

As a final remark to this Section, it is important to highlight that the statistics presented in this Section refer to all municipality in one given year, as they are based on data on local business tax rate of 2015 only. An obvious concern is that, looking at a cross-section only, Twin Tax Differences are a result of tax changes, which will trigger an adjustment of tax rates in neighboring municipalities in the future. In order to investigate the persistence of TTD, I compared data for 2012 with those from 2015, and computed the "persistent TTD", i.e. differences in tax rates that remain stable from 2012 to 2015. The results on the significance of Twin Tax Difference do not change by much, in fact 31% of all municipalities have a persistent Twin Tax Difference of at least two percentage points, 9% a persistent Twin Tax Difference of more than three percentage points.

In sum, this section has presented evidence on the significance of Twin Tax Differences between German municipalities. In a first step, it was shown that Twin Tax Differences of at least two percentage points exist for about 35% of German municipalities. In a second step, it was shown that Twin Tax Differences materialize in many different regions of Germany and are locally concentrated. In the next section, I construct a theoretical model that proposes a structural mechanism behind the emergence of Twin Tax Differences. The model shows how the emergence of Twin Tax Differences depends on parameters of the economic environment.

3 Model

Consider a region consisting of two homogeneous municipalities, A and B. Municipalities autonomously choose their local business tax rates to maximize social welfare within the municipality. There exist two classes of firms in the economy: Regional firms choose where to locate only within the region, foreign firms consider the region as one of their (large) number of alternatives. In order to attract foreign firms to the region, a municipality needs to invest in its infrastructure, e.g. it has to develop an industrial park. The region has a common labor market, which is defined by two characteristics. First, inhabitants of A and B can commute at zero costs between the two municipalities. Second, workers from outside potentially commute into the region, however, their commute is associated with some costs. The model spreads out over three stages. On the first stage, municipalities simultaneously announce whether to develop an industrial park. On the second stage, based on the decision to develop, municipalities simultaneously choose tax rates to maximize social welfare. On the third stage, firms choose their location, pay taxes and workers commute. This section introduces the set-up for households and firms, lines out their actions/objectives and discusses optimal behavior. The next section analyses the implications on tax competition.

3.1 Households

Both municipalities are populated by a continuum n of households. Households inelastically supply one unit of labor, they commute within the region at zero costs, however, they are not able to take a job outside the region. Accordingly, the inelastic labor supply of domestic workers in the region is $2n$ if the wage on the regional labor market w is (weakly) positive, and zero otherwise. Finally, agents generate utility from the consumption of private and public goods: $\text{U} = w + \lambda \frac{G_j}{n}$. Both types of consumption are perfect substitutes, however, the value of public good consumption is weighed by $\lambda \in (0, \infty)$. Finally, $\frac{G_j}{n}$ captures the level of per-capita public good provision in the (fixed) household's municipality of residence.

In addition to the domestic households, there exists a mass n_0 of households outside of the region. A household i from outside supplies one unit of labor inside the region if and only if the difference between the regional and the outside wage, $w - w_0$, is larger than the idiosyncratic commuting cost Δ_c^i. This idiosyncratic commuting costs are uniformly distributed on a closed interval $[0, \Delta_c]$, so that the labor supply from outsiders in the region is $\frac{w-w_0}{\Delta_c} n_0$. Assuming $w_0 \geq 0$, the labor supply in the region is given by:

$$L_S = \begin{cases} 2n + n_0 & \text{if } w \geq w_0 + \Delta_C \\ 2n + \frac{w-w_0}{\Delta_C} n_0 & \text{if } w \in [w_0, w_0 + \Delta_C) \\ 2n & \text{if } w \in (0, w_0) \\ 0 & \text{else} \end{cases} \tag{1}$$

3.2 Firms

The economy hosts two classes of firms, regional and foreign, that is from outside the region. By assumption, if a regional firm chooses to enter the market, it can only locate in A or B. Instead, a foreign firm can choose from a large set of J other municipalities to locate in.

Regional Firms—There exists a mass 1 of heterogeneous regional firms. In general, each firm can develop one project with fixed taxable profits of Π by employing L units of labor[5]. Heterogeneity between regional firms is due to relocation costs. In particular, Δ_R^f captures the monetary costs for firm f from locating in A instead of B. Relocation costs of regional firms are uniformly distributed on an interval $[-\Delta_R, \Delta_R]$. Finally, firms are taxed with respect to an employment based apportionment regime. The share of profits a firm has to pay in a municipality equals its share of employment in that municipality. In total, conditional on developing the project, the location-decision problem of a regional firm f is given by:

$$\max_{L_A, L_B} \Pi^f = \left(1 - \frac{L_A}{L}t_A - \frac{L_B}{L}t_B\right)\Pi - w\left(L_A + L_B\right) - \Delta_R^f L_A$$
$$\text{s.t. } L_A + L_B = L$$

Given the linearity in costs and the apportionment regime, the potential optimality set of a firm is given by $(L_A^*, L_B^*) \in \{(L,0),(0,L),(0,0)\}$, i.e. a firm never locates in both municipalities. All firms choose to develop the project if and only if $\max\{t_A; t_B\} \leq 1$, which will be assumed from now on without loss of generality. Furthermore, it can be shown that a firm locates in A if and only if $\Delta_R^f \leq (t_B - t_A)\,R$, with $R = \frac{\Pi}{L}$. Using these results, the mass of regional firms locating in each of the two municipalities is given by:

$$R_A = \max\left\{0; \frac{\Delta_R + (t_B - t_A)R}{2\Delta_R}\right\} \qquad R_B = \max\left\{0; \frac{\Delta_R + (t_A - t_B)R}{2\Delta_R}\right\}$$

Foreign Firms—There exists a mass 1 of heterogeneous foreign firms. A foreign firm can develop one project with fixed taxable profits of Π by employing L units of labor. Each foreign firm considers location in the region or in a set of municipalities outside of the region. In particular, the set of outside municipalities is heterogeneous across foreign firms. Assuming that each foreign firm prefers a certain municipality m from the set of municipalities outside of the region, we can index foreign firms according to their respective outside option m. Hence, the location problem of a given foreign firm with outside option m can be written as[6]:

$$\max_{L_A, L_m} \Pi^m = \left(1 - \frac{L_A}{L}t_A - \frac{L_m}{L}t_m\right)\Pi - wL_A - w_m L_m$$
$$\text{s.t. } L_A + L_m = L$$

Here, w_m and t_m describe the (given) wage and tax rate of the foreign firm's preferred outside municipality m. Again, due to linearity in costs and the apportionment regime, the potential optimality set of a firm is given by $(L_A^*, L_m^*) \in \{(L,0),(0,L),(0,0)\}$, i.e. a firm never locates in both municipalities. Finally, a firm chooses to locate in municipality A, if and only if $t_A R + w \leq t_m R + w_m$. From now on, assume that the weighted sum of tax rate and wage in the most preferred outside municipality (the RHS of the previous inequality) is uniformly distributed on the interval $[0, \Delta_F]$. Here the best outside option for a firm is assumed to have zero costs, in

[5]The choice to make wage costs non-deductible from the tax base is for tractability. An extension in this direction would complicate the analysis significantly, however, would not change the qualitative results of this project.

[6]For exposition, the location problem is formulated regarding the decision to locate in municipality A.

order to guarantee interior solutions. According to this notation, a foreign firm m locates in A if and only if $\Delta_F^m \geq (t_A R + w)$. Using these results, the mass of foreign firms locating in each of the two municipalities is given by:

$$F_A = \max\left\{0; \frac{\Delta_F - (t_A R + w)}{\Delta_F}\right\} \qquad F_B = \max\left\{0; \frac{\Delta_F - (t_B R + w)}{\Delta_F}\right\}$$

Using the results on foreign and regional firms' behavior, the regional labor demand is given by:

$$L_D = \begin{cases} L & \text{if} \quad \min\{t_A; t_B\}R + w \geq \Delta_F \\ L + F_A & \text{if} \quad t_A R + w < \Delta_F \leq t_B R + w \\ L + F_B & \text{if} \quad t_B R + w < \Delta_F \leq t_A R + w \\ L + F_A + F_B & \text{if} \quad \max\{t_A; t_B\}R + w < \Delta_F \end{cases} \tag{2}$$

3.3 The labor market

This subsection summarizes the findings from the discussion on households' and firms' optimal choice to determine the regional labor market outcomes. The results regarding the regional labor supply are summarized in (1). For the ease of exposition and without loss of generality, fix $w_0 = 0$ for the remainder of the project. The labor supply simplifies to:

$$L_S = \begin{cases} 2n + n_0 & \text{if } w \geq \Delta_C \\ 2n + \frac{w}{\Delta_C} n_0 & \text{if } w \in [0, \Delta_C) \end{cases}$$

The regional labor demand is, instead, summarized by (2). Throughout the paper, the focus is on the upward-sloping part of the labor supply curve. Thus, it has to be guaranteed that $w \leq \Delta_C$. The wage is the highest if the labor demand is maximized, which - in this model - occurs if all foreign firm decide to locate in the region. In this case, labor demand is $L_D = 3L$, which entails the labor market clearing condition $3L = 2n + \frac{w}{\Delta_C} n_0$. Rearranging, we have $w = \frac{3L-2n}{n_0}\Delta_C$. Accordingly, the condition for location on the upward sloping part of the supply curve is: $3L \leq n_0 + 2n$, i.e. the region potentially attracts enough workers to meet the maximum labor demand of firms. Under this assumption, the equilibrium wages are derived as a function of the tax rates only. I start with the case in which none of the municipalities in the region hosts foreign firms. Then, the wage is given by:

$$w = \frac{L - 2n}{n_0} \qquad \text{if} \qquad \min\{t_A; t_B\} \geq \frac{\Delta_F - \frac{L-2n}{n_0}\Delta_C}{R}$$

From here on, for the ease of exposition and without loss of generality, assume that $L = 2n$. Accordingly, in a situation with no foreign firms the wage is given by:

$$w = 0 \qquad \text{if} \qquad \min\{t_A; t_B\} \geq \frac{\Delta_F}{R}$$

Instead, in a situation in which only A attracts foreign firms the wage is given by:

$$w = \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}}(\Delta_F - t_A R) \qquad \text{if} \qquad t_A \leq \frac{\Delta_F}{R}$$

$$t_B \geq \frac{\frac{n_0}{\Delta_C}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}}\frac{\Delta_F}{R} + \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_c} + \frac{L}{\Delta_F}} t_A$$

Similarly, if only B attracts foreign firms the wage is given by:

$$w = \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} (\Delta_F - t_B R) \qquad \text{if} \qquad t_B \leq \frac{\Delta_F}{R}$$

$$t_A \geq \frac{\frac{n_0}{\Delta_C}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} \frac{\Delta_F}{R} + \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_c} + \frac{L}{\Delta_F}} t_B$$

Finally, if both municipalities attract mobile firms, the wage is given by:

$$w = \frac{2\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + 2\frac{L}{\Delta_F}} \left(\Delta_F - \frac{t_A + t_B}{2} R \right) \qquad \text{if} \qquad t_A \leq \frac{\frac{n_0}{\Delta_C}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} \frac{\Delta_F}{R} + \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} t_B$$

$$t_B \leq \frac{\frac{n_0}{\Delta_C}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} \frac{\Delta_F}{R} + \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} t_A$$

3.4 Location of foreign firms

As already discussed in the previous subsection, the foreign firms' decision to locate in the region critically depends on the level of tax rates in both municipalities. This subsection summarizes the results on their location patterns and provides intuition for them.

A necessary condition for a municipality j to host foreign firms is that $t_j \leq \frac{\Delta_F}{R}$. If this condition fails, tax costs in municipality j will be higher than total costs outside the region for all foreign firms. Accordingly, irrespective of the wage level, none of the foreign firms will ever invest in the municipality. If a municipality's tax rate is lower than the threshold $\frac{\Delta_F}{R}$, however, the region's wage level will play a critical role in determining the location pattern of foreign firms.

Consider a situation in which t_B is sufficiently high, so that municipality B will never host any foreign firm. In such a case, $t_A \leq \frac{\Delta_F}{R}$ is a necessary and sufficient condition for municipality A to host a positive mass of foreign firms. If municipality A chooses a tax rate above the threshold, no foreign firm will locate in the region so that $w = 0$. If A charges a tax rate below the threshold, the wage will start to rise above 0, but the sum of wage and tax costs will remain below the outside costs for some foreign firms.

If both municipalities charge a tax rate below $\frac{\Delta_F}{R}$, they will attract foreign firms, not taking into account wage costs. In such a situation, the condition for a municipality's tax rate to attract foreign firms depends on the precise level of the other municipality's tax rate. Consider a situation with low tax rates in A and B, so that both municipalities host some foreign firms. If the tax rate in A increases, less foreign firms will locate in A and so, the regional wage decreases. Accordingly, for a given tax rate in B, the mass of foreign firms locating in B increases, since wage costs are cut. So, the mass of foreign firms locating in B increases in the tax rate of A, due to the interrelated labor markets within a region.

Taking these conditions into account, the 2-dimensional support of tax rates, $[0, 1]^2$, can be split into four different areas, depicting the location pattern of foreign firms: no foreign firms in the region, foreign firms only in A, foreign firms only in B, foreign firms in both municipalities. Figure 4 represents the areas graphically.

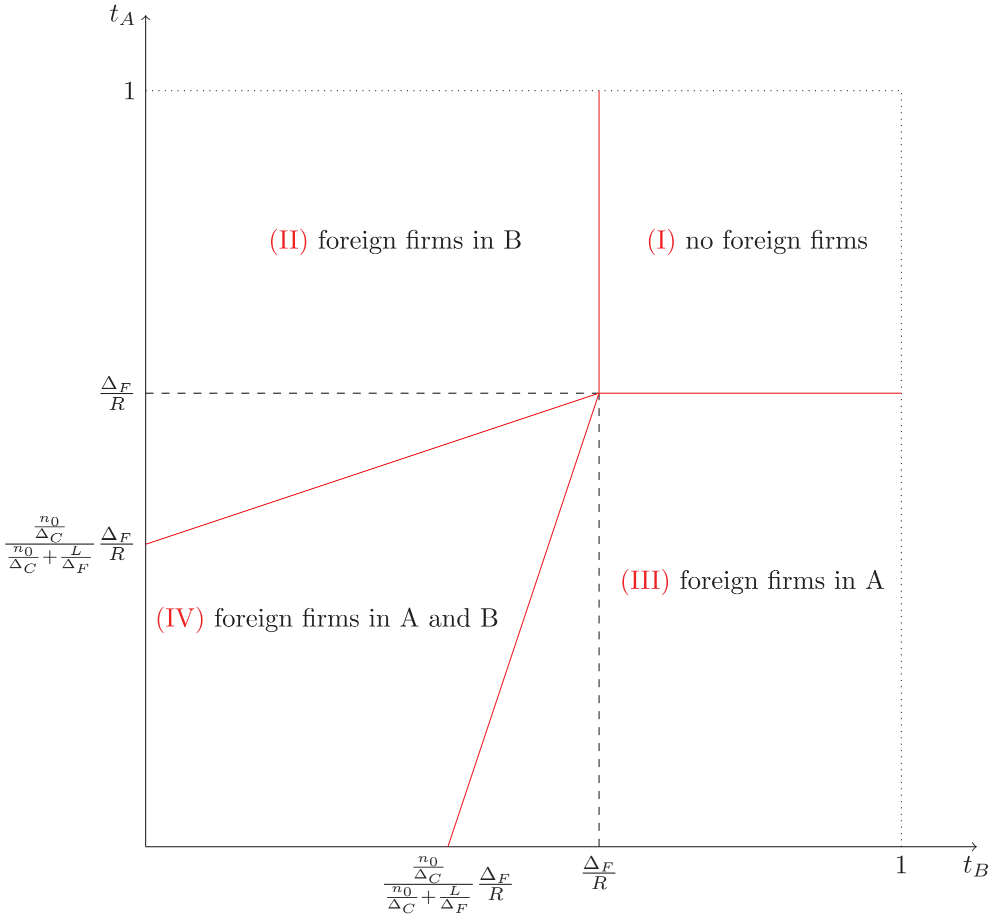


Figure 4: Different location patterns

In Figure 4, the neighbor's tax externality on a municipality's ability to attract foreign firms implies the positive and finite slope of the boundaries of area (IV). If instead the regional wage level was independent of a municipality's tax rate, e.g. due to integration of the regional labor market into a broader labor market ($\Delta_C \to 0$), area (IV) would expand to $\left[0, \frac{\Delta_F}{R}\right]^2$. In such an environment, the ability of a municipality to attract mobile firms would be independent of the neighbor's tax rate.

4 Tax Competition

In the model, tax competition stretches out over two stages. On the first stage, municipalities simultaneously decide to compete for foreign firms. In reality, this decision to compete could involve the development of an industrial park or alike. On the second stage, municipalities simultaneously choose their corporate income tax rates, conditional on their decision to compete for foreign firms[7]. The equilibrium of the model will be determined as a subgame-perfect Nash equilibrium and is solved by backward induction. First I solve for the tax rate equilibrium, given the municipalities' decisions to compete for foreign firms. Second I use these results to discuss if a municipality indeed decides to compete for foreign firms. The sequentiality of tax competition is solely to ease the exposition of our results. In fact, all the qualitative results would carry over to a situation in which municipalities choose tax rates only.

[7] The sequential set-up has previously been used in the literature on local business taxation, e.g. Jayet and Paty (2006)

No foreign firms

We start with constructing a tax rate equilibrium in the subgame in which both municipalities have decided not to compete for foreign firms. In such a situation, the maximization problem of a utilitarian municipality A is given by:

$$\max_{t_A} SWF_A = C_A n + \lambda G_A$$
$$\text{s.t.} \quad C_A = 0, \quad G_A = t_A \left[0.5 + \frac{(t_B - t_A)R}{2\Delta_R} \right] \Pi.$$

Since municipalities have decided not to compete for foreign firms, wages are independent of the tax rates and fixed at $w = 0$. Further, municipalities raise tax revenues only from regional firms. The maximization problem of B is set up symmetrically and the interior[8] equilibrium tax rates are given by:

$$t_A^{NN} = t_B^{NN} = \frac{\Delta_R}{R}$$

Foreign firms in A

This subsection constructs a tax rate-equilibrium in the subgame in which only A has decided to compete for foreign firms. In such a situation, the maximization problem of a utilitarian municipality A is given by:

$$\max_{t_A} SWF_A = C_A n + \lambda G_A$$
$$\text{s.t.} \quad C_A = w = \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} \left(\Delta_F - t_A R \right),$$
$$\text{s.t.} \quad G_A = t_A \left[0.5 + \frac{(t_B - t_A)R}{2\Delta_R} \right] \Pi + t_A \frac{\Delta_F - (t_A R + w)}{\Delta_F} \Pi.$$

Since municipality A has decided to compete for foreign firms, the regional wage depends on the tax rate of A. The smaller t_A, the more foreign firms are attracted to the region, the higher the regional wage. However, the strength of how the tax rate translates into higher wages depends on the parameters of the model. For example, if the regional labor market integrates due to vanishing commuting costs of workers from outside, $\Delta_C \to 0$, regional wages will not react to the inflow of foreign firms. In such a situation, the increase in labor demand is perfectly matched by a higher labor supply from outside workers.

Municipality B has decided not to compete for foreign firms. Accordingly, the maximization problem of B is given by:

$$\max_{t_B} SWF_B = C_B n + \lambda G_B$$
$$\text{s.t.} \quad C_B = w = \frac{\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + \frac{L}{\Delta_F}} \left(\Delta_F - t_A R \right),$$
$$\text{s.t.} \quad G_B = t_B \left[0.5 + \frac{(t_A - t_B)R}{2\Delta_R} \right] \Pi.$$

Since municipality B has decided not to compete for foreign firms, the regional wage does not depend on the tax rate of B. Furthermore, B generates tax revenues only from regional firms.

[8] The conditions on interiority and potential corner solutions are discussed in Appendix A.

In an interior asymmetric Nash-equilibrium[9], tax rates are given by:

$$t_A^{NF} = \frac{0.75 + \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C}+\frac{L}{\Delta_F}} - \frac{\frac{L}{\Delta_F}}{\frac{L}{\Delta_F}+\frac{n_O}{\Delta_C}}\frac{n}{\lambda}\frac{R}{\Pi}}{0.75\frac{R}{\Delta_R} + 2\frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C}+\frac{L}{\Delta_F}}\frac{R}{\Delta_F}} \qquad t_B^{NF} = \frac{0.75 + 0.5\left(1+2\frac{\Delta_R}{\Delta_F}\right)\frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C}+\frac{L}{\Delta_F}} - 0.5\frac{\frac{L}{\Delta_F}}{\frac{L}{\Delta_F}+\frac{n_O}{\Delta_C}}\frac{n}{\lambda}\frac{R}{\Pi}}{0.75\frac{R}{\Delta_R} + 2\frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C}+\frac{L}{\Delta_F}}\frac{R}{\Delta_F}} \tag{3}$$

The tax rates of the two municipalities differ due to two distinct reasons.

On the one hand, the tax rate-elasticities of foreign and regional tax base potentially differ. Hence, because of a standard inverse elasticity-rule, a municipality aiming to generate tax revenues from foreign firms will choose a different tax rate. In a 'no foreign firms'-equilibrium the elasticity of the regional tax base is given by $\epsilon_{R_A,t_A} = -\frac{R}{\Delta_R}$, while the elasticity of the foreign tax base would be given by $\epsilon_{F_A,t_A} = -\frac{R}{\Delta_F - \Delta_R}$. Accordingly, from a revenue-maximizing perspective A would choose a tax rate lower than that in the 'no foreign firms'-equilibrium if $\epsilon_{R_A,t_A} < \epsilon_{F_A,t_A}$, i.e. $\Delta_F < 2\Delta_R$.

On the other hand, if a municipality competes for foreign firms, its tax rate will have an impact on the regional wage as well. In fact, once the municipality positively values the regional wage level in its objective, it will choose a tax rate on the upward-sloping part of the (constrained) Laffer-curve. In the model, the relative weight on private consumption decreases in λ, implying that - as such parameter increases - the equilibrium tax rate increases, while the asymmetry goes to zero.

Finally, for $\lambda \to \infty$ and $\Delta_F \to 2\Delta_R$ the asymmetric tax rates converge to the 'no foreign firms'-equilibrium tax rates. In such case, foreign and regional tax base are symmetric from a municipality's point of view and so, tax rates do not depend on the kind of tax base it is competing for.

Foreign firms in A and B

This subsection constructs a tax rate equilibrium in the subgame in which both municipalities have decided to host foreign firms. In such a situation, the maximization problem of a utilitarian municipality A is:

$$\begin{aligned}
&\max_{t_A} SWF_A = C_A n + \lambda G_A \\
&\text{s.t.} \quad C_A = w = \frac{2\frac{L}{\Delta_F}}{\frac{n_0}{\Delta_C} + 2\frac{L}{\Delta_F}}\left(\Delta_F - \frac{t_A + t_B}{2}R\right), \\
&\text{s.t.} \quad G_A = t_A\left[0.5 + \frac{(t_B - t_A)R}{2\Delta_R}\right]\Pi + t_A\frac{\Delta_F - (t_A R + w)}{\Delta_F}\Pi.
\end{aligned}$$

Since both municipalities have decided to compete for foreign firms, the regional wage depends on both of their tax rates. The transmission intensity, again, depends on the parameters of the model, e.g. the commuting costs Δ_C. The maximization problem of B is set up symmetrically and, in an interior Nash-equilibrium[10], tax rates are:

$$t_A^{FF} = t_B^{FF} = \frac{0.5 + \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C}+2\frac{L}{\Delta_F}} - \frac{\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C}+2\frac{L}{\Delta_F}}\frac{R}{\Pi}\frac{n}{\lambda}}{\frac{R}{2\Delta_R} + \frac{R}{\Delta_F}\frac{2\frac{n_O}{\Delta_C}+\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C}+2\frac{L}{\Delta_F}}} \tag{4}$$

The symmetric tax rates differ from a 'no foreign firms'-equilibrium due to two different reasons.

[9] The conditions on interiority and potential corner solutions are discussed in Appendix A.
[10] The conditions on interiority and potential corner solutions are discussed in Appendix A.

On the one hand, tax rates are structurally lower due to the incentives to attract foreign firms in order to increase private consumption. If municipalities put a positive and finite weight λ on private consumption they will choose a tax rate on the upward sloping part of the (constrained) Laffer curve. However, if the weight on private consumption decreases, i.e. λ increases, tax rates will also increase to generate higher tax revenues from foreign and regional firms.

On the other hand, the tax rate-elasticities of foreign and regional tax base could also differ. By the same inverse elasticity argument made before, tax rate would differ as a consequence of tax revenue maximization. In fact, it can be shown that tax rates would be lower as compared to a 'no foreign firms'-equilibrium if $\Delta_F < \Delta_R \left[2 + \frac{\frac{L}{\Delta_F}}{\frac{nQ}{\Delta_C}}\right]$. The threshold is comparable to the one derived in the case in which only one municipality competes for foreign firms, however, it now takes into account the existence of a labor market externality. If the externality were not existing, e.g. $\Delta_C \to 0$, both thresholds would align. Otherwise, tax rates would be lower than in the single competition case to counterbalance the higher regional wages.

Finally, for $\lambda \to \infty$ and $\Delta_F \to \Delta_R \left[2 + \frac{\frac{L}{\Delta_F}}{\frac{nQ}{\Delta_C}}\right]$, the 'foreign firms'-equilibrium tax rates converge to the 'no foreign firms'-equilibrium tax rates. In such case, foreign and regional tax base are symmetric from a municipality's point of view and, so, tax rates do not depend on the kind of tax base a municipality is competing for.

Decision to compete

On Stage 1 of the tax competition game, municipalities decide whether or not to compete for foreign firms. In their decision, municipalities take into account the Nash-equilibria in the subgames on Stage 2. The structure of the tax competition game on Stage 1 can be displayed in normal form as follows:

		B: Compete	B: Not Compete
A	Compete	$SWF_A\left(t_A^{FF}, t_B^{FF}\right)$, $SWF_B\left(t_A^{FF}, t_B^{FF}\right)$	$SWF_A\left(t_A^{FN}, t_B^{FN}\right)$, $SWF_B\left(t_A^{FN}, t_B^{FN}\right)$
A	Not Compete	$SWF_A\left(t_A^{NF}, t_B^{NF}\right)$, $SWF_B\left(t_A^{NF}, t_B^{NF}\right)$	$SWF_A\left(t_A^{NN}, t_B^{NN}\right)$, $SWF_B\left(t_A^{NN}, t_B^{NN}\right)$

Figure 5: Decision to compete

On Stage 1, there are four candidates for a subgame-perfect Nash equilibrium in pure strategies: 'Compete-Compete', 'Compete-Not Compete', 'Not Compete-Compete', 'Not Compete-Not Compete'. Given the discussion in the previous subsections, a situation with Twin Tax Differences is defined as follows:

Definition 1. *Tax competition between adjacent municipalities generates a situation with Twin Tax Differences in the region if and only if in a subgame-perfect Nash equilibrium the equilibrium strategies on Stage 1,* $\{\mathrm{S}_A; \mathrm{S}_B\}$*, are either* {'Compete';'Not Compete'} *or* {'Not Compete';'Compete'}.

5 Results

This section introduces and explains the main results of the project. Initially, the subgame-perfect Nash equilibrium of the tax competition game is introduced, showing how the type of equilibrium depends on the main parameters of the model. It follows a discussion of the main mechanisms underlying the results. Furthermore, this section provides a discussion on whether and under which circumstances the decentralized equilibrium maximizes regional welfare. The last part sheds light on the results' policy implications, both in terms of incentive provision and redistribution between municipalities.[11]

5.1 Equilibrium

The equilibrium of the model can be of three different types: symmetric high tax rates without foreign firms in the region, symmetric low tax rates with foreign firms in both municipalities, asymmetric tax rates with only one municipality hosting foreign firms. Within the frame of this paper, the focus is on the emergence of equilibria with asymmetric tax rates, Twin Tax Differences.

The economy is governed by two main parameters of interest. On the one hand, the foreign firms' outside options are parametrized by Δ_F. The higher Δ_F, the easier it is for municipalities of the region to attract foreign firms. In reality, Δ_F could be thought of as the quality of regional infrastructure, like the access to transportation or fast internet. On the other hand, the intensity of commuting streams to a region is parametrized by Δ_C. The higher Δ_C, the smaller the amount of workers commuting into a region, for a given difference in wages. In reality, Δ_C could be thought of as driving proximity to a metropolitan area. Based on the level of those two parameters, different sets of equilibria emerge.

Proposition 1. *There exists an intensity of in-commuting, Δ_C^{thr}, such that for*

- $\Delta < \Delta_C^{thr}$ *Twin Tax Differences do not emerge in a decentralized equilibrium.*
- $\Delta \geq \Delta_C^{thr}$ *Twin Tax Differences can emerge in a decentralized equilibrium.*

For sufficiently small levels of Δ_C, Twin Tax Differences do not emerge in a decentralized equilibrium. If the commuting costs of workers into the region are too small, any attraction of foreign firms leads to an inflow of workers to the region such that the equilibrium wage on the regional labor market remains mostly unaffected. The regional labor market has integrated in a broader labor market with a given wage level of w_O. In this case, the only remaining interdependence between municipalities is the regional tax base externality. Such externality gives rise only to symmetric equilibria, either with high or low tax rates, but never allows for the existence of an asymmetric equilibrium.

As Δ_C increases, the regional labor market separates from the rest of the world. Accordingly, an inflow of foreign firms to the region increases wages on the regional labor market, as the increase in labor demand is not fully offset by an increase in labor supply due to commuters. In such situations, the labor market tax

[11] The presentation of the results focuses on a parameter space, so that the attraction of foreign firms requires a lower level of taxation than it would realize in a 'Not Compete-Not Compete'-equilibrium. Otherwise, the attraction of foreign firms would be 'free lunch'. Finally, under such restriction, the results of the sequential competition model align with a model in which municipalities choose tax rates only.

externality materializes. If a municipality attracts foreign firms by lowering the tax rate, this directly raises a neighbor's welfare due to an increase in regional wages and so private consumption. At the same time, it diminishes the ability of a neighbor to attract foreign firms, as the labor costs on the regional market increase. Under the presence of a significant labor market externality, Twin Tax Differences exist in equilibrium.

In such equilibrium, one municipality charges a low tax rate and attracts foreign firms to the region, while the other municipality charges a high tax rate and only hosts regional firms. Under the presence of a sizable labor market externality, no municipality has an incentive to deviate from an asymmetric equilibrium. On the one hand, the municipality with the low tax rate attracts a larger tax base of regional *and* foreign firms. In addition, given the attraction of foreign firms, the municipality's low tax rate increases the wage in the region. On the other hand, the municipality with the high tax rate only hosts a smaller share of regional firms and no foreign firms. However, due to the presence of the labor market externality, the positive foreign firm tax base effect from lowering the tax rate is attenuated for the high tax municipality. Given that labor costs are already high on the regional labor market, a lowering of the tax rate does not lure much additional foreign firms' investment to the municipality.

Proposition 2. *For every level of $\Delta_C \geq \Delta_C^{thr}$, there exists a $\Delta_F^{up}(\Delta_C)$ and a $\Delta_F^{low}(\Delta_C)$ so that Twin Tax Differences emerge in a decentralized equilibrium if and only if $\Delta_F \in \left[\Delta_F^{low}(\Delta_C), \Delta_F^{up}(\Delta_C)\right]$.*

Twin Tax Differences do not generally exist for high levels of Δ_C, but their existence additionally depends on the level of Δ_F. In particular, Twin Tax Differences only exist for intermediate levels of Δ_F. If Δ_F is too high, the high tax municipality has an incentive to deviate, given that the municipality is still able to attract enough foreign firms, despite the higher labor costs. On the contrary, if Δ_F is too low, it will be the municipality with the low tax rate which has an incentive to increase its tax rate, given that the positive foreign tax base effect is not large enough. Figure 6 summarizes these findings graphically as a result of a numerical simulation of the model.

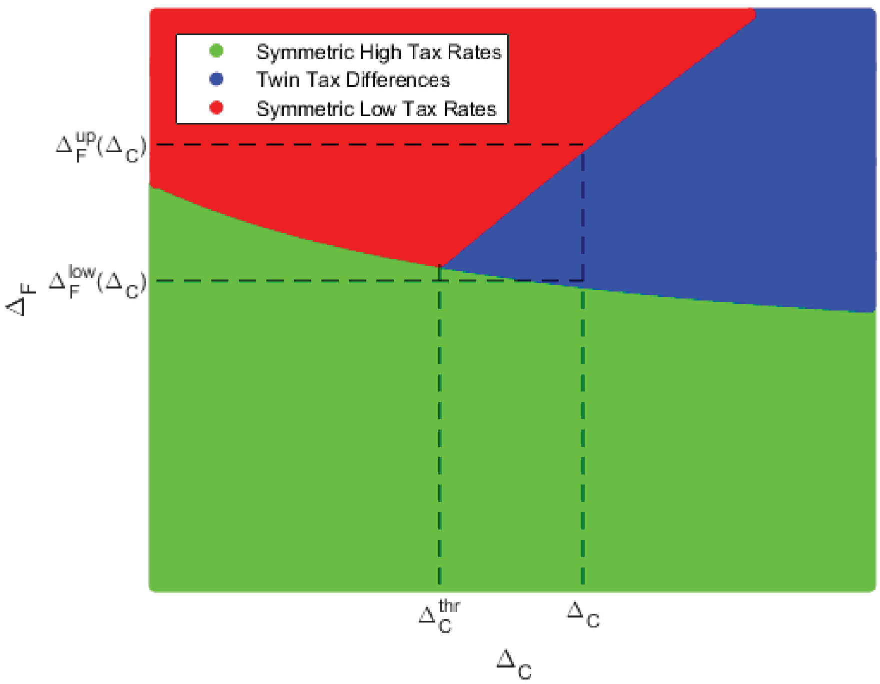


Figure 6: Equilibria of the Tax Competition

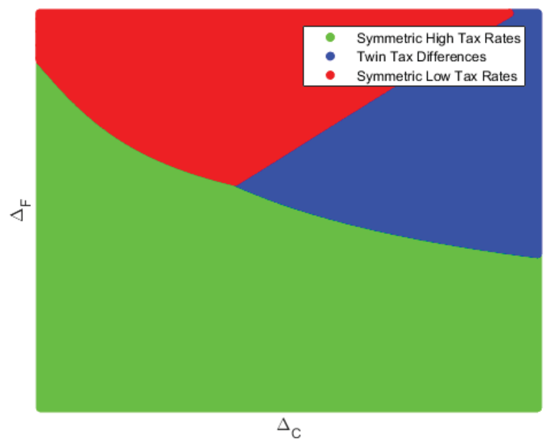
In summary, this subsection has lined out the conditions under which Twin Tax Differences can emerge in a decentralized tax competition equilibrium. It was shown that an equilibrium with Twin Tax Differences exists only if the labor market externality of the tax policy of an adjacent municipality is sufficiently negative, e.g. due to the existence of a common labor market. However, the ability of such externality to induce Twin Tax Differences hinges on the general competitive position of the region under consideration. If the region is highly attractive for investment, the labor market externality does not suffice to prevent all municipalities from choosing low tax policies, if the region is not attractive for investment on the first place, no municipality will actually choose a low tax policy to trigger the externality.

5.2 Regional Welfare

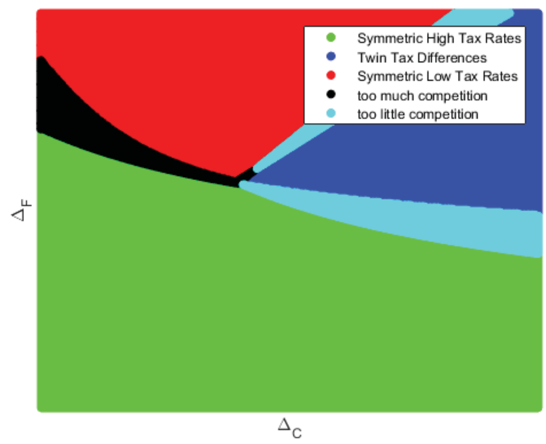
This subsection discusses the consequences of decentralized taxing rights on regional welfare, defined as the sum of municipalities' social welfare. In particular, it compares welfare levels at the commitment stage, i.e. at the point when municipalities non-cooperatively decide whether to compete for foreign firms. The point of this exercise is to discuss whether, from a regional viewpoint, decentralized taxing rights imply too much or too little tax competition, as measured by the number of municipalities competing for foreign firms, which is similar to the idea of Matsumoto (2010).

The results are tightly related to the two externalities. As standard in tax competition models, the existence of a tax base externality implies a tendency of excessive competition in equilibrium. In their objective, municipalities do not take into account the positive tax base effect of a higher tax rate on the neighbor's tax base and, therefore, the level of taxation is too low from a regional viewpoint. The effect of the labor market externality on the efficiency property of the equilibrium is just reversed. In fact, in choosing its tax rate, a municipality does not take into account the negative wage effect of a higher tax rate on the neighbor's private consumption.

Accordingly, due to this labor market externality, tax rates tend to be excessively high in equilibrium from a regional viewpoint.



(a) Welfare-maximizing



(b) Inefficiency

Figure 7: Regional welfare

Panel (a) of Figure 7 presents the competitive situations that maximize regional welfare. The general structure of the regional welfare-maximizing decisions to compete is similar to the results on the decentralized equilibria, discussed in Subsection 5.1: Only if the labor market externality becomes sufficiently strong, a Twin Tax Difference can maximize regional welfare.

The implications of decentralized taxing rights on regional welfare are depicted in Panel (b) of Figure 7. The results depend on the relative strength of the two externalities. If the labor market externality becomes relatively strong compared to the tax base externality, the equilibrium is more likely to incorporate too little competition for foreign firms and vice versa. In terms of the model's parameters, this means that, for sufficiently small levels of Δ_C, there exists a set of intermediate levels of Δ_F such that the equilibrium implies too much competition from a regional viewpoint. As Δ_C increases, however, the set of Δ_F that induces such a situation shrinks. Finally, there exists a level of Δ_C above which no equilibria exist anymore, in which there is too much competition from a regional viewpoint.

By contrast, for relatively high levels of Δ_C there exist intermediate levels of Δ_F that would imply too little competition in equilibrium from a regional viewpoint. The reason is that the labor market externality is sufficiently strong, so that the commuting stream is not enough anymore to balance out wages on the regional labor market. Thus, municipalities should compete even more to increase regional wages and, so, private consumption in the whole region.

Finally, the definition of regional welfare neglects moving costs of regional firms. In fact, Twin Tax Differences trigger tax-motivated relocation for some regional firms. One could argue that such moving costs constitute a tax-induced inefficiency and should thus be incorporated in the notion of regional welfare. However, it can be shown that taking into account moving costs does not change the qualitative results of this subsection.

In summary, this subsection has shown that Twin Tax Differences can maximize regional welfare. On the one hand, hosting a low tax municipality triggers investment from foreign firms to the region and so increases regional wages. On the other hand, hosting a high tax municipality increases the region-wide tax revenues from regional firms and so mitigates the opportunity tax costs of attracting investment from foregin firms. Twin Tax Differences can emerge as an equilibrium with decentralized taxing rights, as we have seen in the data as well as shown in our theoretical model. However, this subsection additionally identifies situations, in which Twin Tax Differences do not emerge in equilibrium, despite they would maximize welfare of the whole region. This observation is picked up in the discussion on policy implications in the next subsection.

5.3 Policy Implication - Incentives

In the policy debate, tax competition between municipalities is commonly identified as a situation of beggar-thy-neighbor. This perception is derived from the existence of a sizable tax base-externality. Regional firms are the most important source of municipalities' tax revenues and, hence, tax competition for such revenues is fierce. As a consequence, interventions in tax competition between municipalities were designed to stabilize symmetric high tax equilibria.

In this paper, this common intuition unambiguously holds true only for situations in which the labor market externality is sufficiently weak. For a low level of Δ_C the model generates symmetric equilibria in high or low tax rates only. Moreover, for some intermediate levels of Δ_M there is too much tax competition from the perspective of regional welfare. In this case, an intervention of a regional body to correct for the tax base externality performs well in improving regional welfare by providing disincentives against excessive competition.

As the labor market externality gains importance, the adequacy of such policies changes. For a sufficiently large Δ_C, the region should rather foster tax competition of municipalities. The reason is that a municipality does not internalize the positive wage effect of a tax rate decrease on the level of private consumption in the neighboring municipality.

In Germany, regions particularly support cooperation of municipalities in the development of industrial parks, inducing them to split development costs and tax revenues. These sharing agreements are implemented to incentivize competition for foreign firms without depressing the level of taxation in the region too much. In the presence of a sizable labor market externality, however, my model suggests that such policies could actually be detrimental for regional welfare. Instead, regions could gain from having one municipality more aggressively competing for foreign firms[12].

5.4 Policy Implication - Redistribution

In a symmetric equilibrium, both municipalities generate the same social welfare, while in an equilibrium with Twin Tax Differences, the levels of social welfare diverge. Which of the municipalities is better off in a situation with Twin Tax Differences is indeterminate.

In an equilibrium with Twin Tax Differences one of the municipalities chooses a low tax rate to attract foreign firms to the region, while the other one charges a higher tax rate and hosts regional firms only. Both municipalities equally gain from the inflow of foreign firms through an increase in private consumption. So, the differences in social welfare are completely due to differences in tax revenues.

[12] From the viewpoint of the region, the line of argumentation goes along the insights provided by Hong and Smart (2010).

Which of the municipalities generates higher tax revenues is indeterminate. On the one hand, the low tax municipality broadens its tax base due to the inflow of foreign firms and a higher share of regional firms. On the other hand, the high tax municipality keeps its tax rate closer to the revenue-maximizing rate for regional firms. Which of the effects, broader base or higher rate, generates higher tax revenues depends on the level of the underlying parameters.

For a given tax rate, the mass of foreign firms locating in the low tax municipality increases in Δ_F and decreases in Δ_C. Thus, the social welfare of the low tax municipality mechanically increases in Δ_F and decreases in Δ_C. As the foreign tax base increases, the strategic response of a low tax municipality is to increase its tax rate, because the tax rate effect gains importance. In response, the high tax municipality will increase its tax rate as well, due to the strategic complementarity triggered by the tax base externality through the regional tax base.

Overall, tax revenues of the low tax municipality increase by more if Δ_F increases or Δ_C decreases. The intuition is that a change in these parameters mechanically affects only tax revenues of the low tax municipality, while changes in the tax revenues of the high tax municipality are purely strategic. In consequence, the set of asymmetric equilibria can be split up in two parts, in accordance to the ranking of social welfare in the municipalities. In an equilibrium with Twin Tax Differences, the low tax municipalities generates a higher social welfare if the ability to attract foreign firms, given by Δ_F, is sufficiently high and the intensity of the labor market externality, given by Δ_C, is respectively weak.

The results have policy relevance, if you consider a regional government who wants to harmonize social welfare across its municipalities. In particular, the results show that the level of a municipality's tax rate does not suffice to determine the direction of redistribution. There are certain cases in which the low tax municipality should be refunded for its effort to increase private consumption in the region. However, in other instances the low tax municipality attracts enough foreign firms and rather, public funds should be channelled to the high tax municipality as a compensation for the decrease in regional tax revenues.

6 Empirical Evidence

This section provides some evidence on the main parameters of the model. Once again, the setting of German local business taxation is used as a testing ground.

6.1 Labor Market Externality

A necessary condition for the emergence of Twin Tax Differences is the existence of a significant labor market externality. In equilibrium, the low tax municipality faces a higher labor demand from foreign and regional firms. A labor market externality exists if the higher labor demand is not fully offset by a higher labor supply of workers from the low tax municipality or from outside the region, but triggers a commuting flow from the high tax to the low tax municipality.

In order to find evidence for the existence of a labor market externality, I collect data from the regional data base of the German statistical offices (*Regionaldatenbank*). The data does not contain information on bilateral commuting flows. However, on the municipality level the data provides information on the persons employed by place of residence and by place of employment. Accordingly, the in-commuting saldo for each municipality is computed as the difference of the number of people employed by place of employment minus those persons employed by place of residence.

The intuition is that a municipality with a positive Twin Tax Difference exports the creation of private consumption to its low tax neighbor, by the mean of a large degree of out-commuting. In the data, such a pattern would imply the in-commuting saldo to decrease in the Twin Tax Difference. The evidence is summarized in Figure 8.

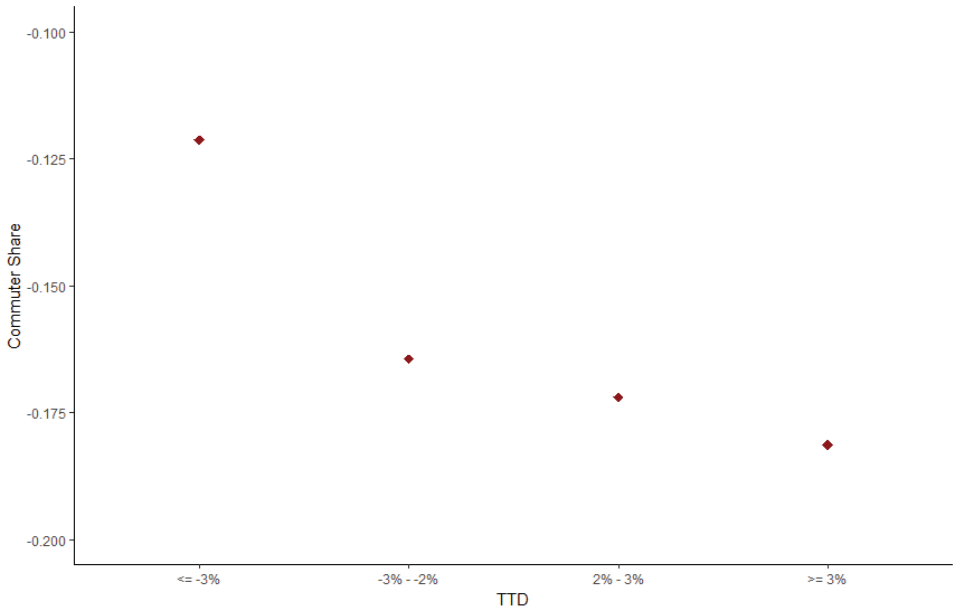


Figure 8: Twin Tax Differences and commuter share

The commuter share is computed as the ratio of a municipality's in-commuting saldo relative to its population. Figure 8 presents the average commuter shares, grouped by the existence of a certain degree of Twin Tax Difference. Generally, the commuting share is negative across all Twin Tax Differences. Most commuting is observed from villages to big cities. Since big cities rarely have a neighbor of the same size in the same region, our sample contains very few big cities and so, the average commuter share is negative throughout.

Figure 8 shows a negative relation between the commuter share and the Twin Tax Difference. In numbers, for municipalities at opposite sides of a Twin Tax Difference wider than three percentage points, the Commuter Share differs by six percentage points on average. This evidence speaks in favor of a labor market externality, so that, in a situation with Twin Tax Differences, high tax municipalities export the creation of private consumption to their low tax neighbors.

6.2 Foreign Firms

The ability of a region to attract foreign firms (firms from outside the region) is the second necessary condition for the existence of a tax competition equilibrium with Twin Tax Differences. If a region does not have the ability to attract foreign firms, municipalities compete for regional firms only and the tax competition equilibrium will necessarily be symmetric.

In order to provide evidence on the availability of a foreign firms' tax base, I once again use data from the regional data base of the German statistical offices. The database contains information on the migration flows of firms on the regional (*Kreis*) level. I construct a dataset on the regional level, using the highest absolute Twin Tax Difference within a region as a variable. Finally, the sample of regions is split by the size of the Twin Tax Difference and the average number of firms newly arriving to the region between 2012 and 2015 is computed within every sub-sample. Figure 9 presents the results of this exercise.

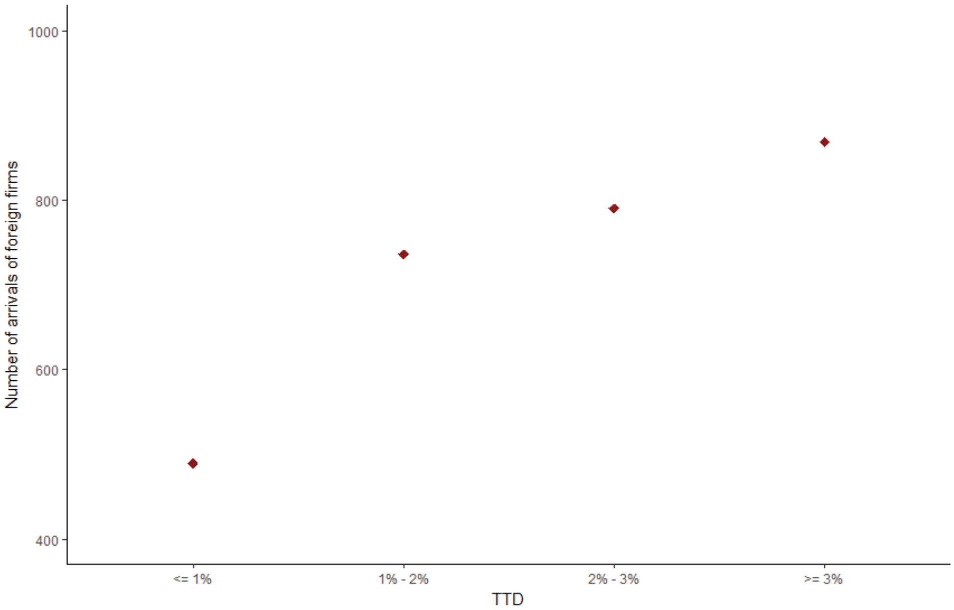


Figure 9: Twin Tax Differences and the arrival of foreign firms

The average number of newly arriving firms increases in the size of a region's Twin Tax Difference. There are two possible interpretations for this result. First, it might be the case (as it can additionally be shown) that regions with large Twin Tax Differences tend to have lower minimal local business tax rates. In this light, the result on newly arriving firms would provide evidence that municipalities are able to attract firms from outside the region by choosing a low tax policy, in line with the findings of Janeba and Osterloh (2013). Second, the result could be seen as evidence that, indeed, neighboring municipalities will find it optimal to tolerate a positive Twin Tax Difference if the low tax neighbor is able to attract firms from outside the region, in line with this paper's theoretical argument.

7 Conclusion

This paper sets out for highlighting an empirical pattern in local business tax rates, which has so far been largely overlooked by the literature on local business taxation: the emergence of Twin Tax Differences.

Using data on local business tax rates in Germany, the empirical significance of Twin Tax Differences has been documented. Controlling for location and size, more than 35 % of the German municipalities are located in areas with Twin Tax Differences of at least two percentage points. Furthermore, it was also shown how Twin Tax Differences are not a regional phenomenon, but can be observed in different parts of Germany.

In order to understand the conditions under which Twin Tax Differences are likely to emerge, the paper introduces a theoretical model that generates Twin Tax Differences as an equilibrium outcome of tax competition. In difference to previous models on local business tax competition, our model generates Twin Tax Differences even in a situation with fully homogeneous municipalities. The key to this result is the presence of a significant labor market externality, so that a municipality's decision to aggressively compete for tax revenues deteriorates the competitive position of its neighbors.

The externality arises only if the region under consideration incorporates certain features. First, the regional labor market needs to be partly segregated. Second, the region needs to be able to attract investment from outside, while still maintaining a regional tax base of a certain importance. The last part of the paper provides initial evidence on those characteristics.

The main purpose of this project so far is to raise the literature's awareness of Twin Tax Differences. Future work is needed to strengthen the theoretical part by incorporating other important features of local business taxation, for example the existence of vertical externalities. Finally, the evidence presented on the model's main parameters is so far only a sketch. Much more work has to be done here in order to provide clean evidence for the parameters' existence.

A Tax rate equilibrium - corner solutions

no foreign firms

For an interior solution, it has to hold that $t_A^{NN} = \frac{\Delta_R}{R} \geq \frac{\Delta_M}{R}$. If this relation fails to hold true, the corner solution is given by $t_A^{NN} = t_B^{NN} = \frac{\Delta_M}{R}$.

foreign firms only in A

For an interior solution, the tax rates have to be within area (III) of Figure 1. However, for the model's parameter space, a location in (III) does not hold unambiguously true. Given that $t_A^{NF} \leq t_B^{NF}$, equilibria can potentially locate in (I), (III), (IV) or even with negative tax rates in A.

If the solution is in area (III), the solution will be interior. If the solution is in area (I), the corner solution will be:

$$t_A^{NF} = \frac{\Delta_F}{R}$$
$$t_B^{NF} = 0.5\left[\frac{\Delta_R}{R} + \frac{\Delta_F}{R}\right]$$

Such an equilibrium will emerge, if it holds true that:

$$\frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} + \frac{\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} \frac{n}{\lambda} \frac{R}{\Pi} \leq 0.75\left(1 - \frac{\Delta_F}{\Delta_R}\right)$$

If the solution has negative tax rates of A, the corner solution will be:

$$t_A^{NF} = 0$$
$$t_B^{NF} = 0.5 \frac{\Delta_R}{R}$$

Such an equilibrium will emerge, if it holds true that:

$$\lambda < \frac{\frac{\frac{L}{\Delta_F}}{\frac{L}{\Delta_F} + \frac{n_O}{\Delta_C}} n \frac{R}{\Pi}}{0.75 + \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}}}$$

If the solution is in area (IV), the corner solution will be:

$$t_A^{NF} = \frac{0.5 + \left(1 + 0.5\frac{\Delta_F}{\Delta_R}\right) \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} - \frac{\frac{L}{\Delta_F}}{\frac{L}{\Delta_F} + \frac{n_O}{\Delta_C}} \frac{n}{\lambda} \frac{R}{\Pi}}{\frac{\frac{n_O}{\Delta_C} + 0.5\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} \frac{R}{\Delta_R} + \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + 2\frac{L}{\Delta_F}} \frac{R}{\Delta_F}}$$

$$t_B^{NF} = \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} \frac{\Delta_F}{R} + \frac{\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} \frac{0.5 + \left(1 + 0.5\frac{\Delta_F}{\Delta_R}\right) \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} - \frac{\frac{L}{\Delta_F}}{\frac{L}{\Delta_F} + \frac{n_O}{\Delta_C}} \frac{n}{\lambda} \frac{R}{\Pi}}{\frac{\frac{n_O}{\Delta_C} + 0.5\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} \frac{R}{\Delta_R} + \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + 2\frac{L}{\Delta_F}} \frac{R}{\Delta_F}}$$

Such an equilibrium will emerge, if it holds true that:

$$0.5\frac{\Delta_R}{R} - \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} \frac{\Delta_F}{R} \leq \left[\frac{\frac{L}{\Delta_F}}{\frac{L}{\Delta_F} + \frac{n_O}{\Delta_C}} - 0.5\right] \frac{0.5 + \left(1 + 0.5\frac{\Delta_F}{\Delta_R}\right) \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} - \frac{\frac{L}{\Delta_F}}{\frac{L}{\Delta_F} + \frac{n_O}{\Delta_C}} \frac{n}{\lambda} \frac{R}{\Pi}}{\frac{\frac{n_O}{\Delta_C} + 0.5\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}} \frac{R}{\Delta_R} + \frac{\frac{n_O}{\Delta_C}}{\frac{n_O}{\Delta_C} + 2\frac{L}{\Delta_F}} \frac{R}{\Delta_F}}$$

foreign firms in A and B

For an interior solution, the tax rates have to be such that $t_A^{FF} \in \left[0, \frac{\Delta_F}{R}\right]$.

If $t_A^{FF} < 0$, the corner solution is $t_A^{FF} = t_B^{FF} = 0$. The condition for such a corner solution to emerge is:

$$\lambda \leq \frac{\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + 2\frac{L}{\Delta_F}} \frac{R}{\Pi} n$$

If $t_A^{FF} > \frac{\Delta_F}{R}$, the corner solution is $t_A^{FF} = t_B^{FF} = \frac{\Delta_F}{R}$. The condition for such a corner solution to emerge is:

$$\frac{\frac{1.5\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + 2\frac{L}{\Delta_F}} - \frac{\frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + 2\frac{L}{\Delta_F}} \frac{R}{\Pi} \frac{n}{\lambda}}{\frac{2\frac{n_O}{\Delta_C} + \frac{L}{\Delta_F}}{\frac{n_O}{\Delta_C} + 2\frac{L}{\Delta_F}} + \frac{R}{2\Delta_R} \frac{\Delta_F}{R}} \geq 1$$

References

David R Agrawal, William F Fox, and Joel Slemrod. Competition and subnational governments: tax competition, competition in urban areas, and education competition. *National Tax Journal*, 68(3):701–735, 2015.

Tobias Bohm, Nadine Riedel, and Martin Simmler. Large and influential: firm size and governments' corporate tax rate choice? *Centre for Business Taxation Working Paper*, 2016.

Ralph M Braid. Symmetric tax competition with multiple jurisdictions in each metropolitan area. *The American Economic Review*, 86(5):1279–1290, 1996.

Marius Brülhart, Mario Jametti, and Kurt Schmidheiny. Do agglomeration economies reduce the sensitivity of firm location to tax differentials? *The Economic Journal*, 122(563):1069–1093, 2012.

Xavier Giroud and Joshua Rauh. State taxation and the reallocation of business activity: Evidence from establishment-level data. *Journal of Political Economy*, 127(3):000–000, 2019.

Qing Hong and Michael Smart. In praise of tax havens: International tax planning and foreign direct investment. *European Economic Review*, 54(1):82–95, 2010.

Eckhard Janeba and Steffen Osterloh. Tax and the city—a theory of local tax competition. *Journal of Public Economics*, 106:89–100, 2013.

Hubert Jayet and Sonia Paty. Capital indivisibility and tax competition: Are there too many business areas when some of them are empty? *Journal of Urban Economics*, 60(3):399–417, 2006.

Michael J Keen and Christos Kotsogiannis. Does federalism lead to excessively high taxes? *The American Economic Review*, 92(1):363–370, 2002.

Mutsumi Matsumoto. Entry in tax competition: a note. *International Tax and Public Finance*, 17(6):627–639, 2010.

Ferdinando Monte, Stephen J Redding, and Esteban Rossi-Hansberg. Commuting, migration, and local employment elasticities. *The American Economic Review*, 108(12):3855–90, 2018.

Charles M Tiebout. A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424, 1956.