CESifo WORKING PAPERS

7940
2019

Original Version: November 2019
This Version: July 2020

Tax Systems and Development

Mohammed Mardan

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7940*
Category 1: Public Finance

Tax Systems and Development

Abstract

This paper analyzes the optimal tax system consisting of the tax rate and the capacity of the tax administration. I challenge the conventional wisdom that tax rates and tax administrations' capacity are complements showing that their relationship depends on a country's level of development. Accordingly, developed and developing countries should set higher tax rates and administrative capacities than moderately developed countries. The discrepancy compared to observed practices regarding administrative capacities can be resolved by an aspect previously disregarded in the tax systems literature: the degree of tax administrations' autonomy. With autonomous tax administrations, capacity increases monotonically with development as suggested by the empirical literature. Finally, I analyze how fighting tax havens affects non-haven countries' tax revenues. While developed countries benefit, the same applies to developing countries only if tax administrators and lawmakers tightly coordinate their actions.

JEL-Codes: H250, O230, F230.

Keywords: developing countries, profit shifting, tax administration, tax competition, tax haven.

Mohammed Mardan
Norwegian School of Economics
Department of Business & Management Science
Helleveien 30
Norway – 5045 Bergen
mohammed.mardan@nhh.no

Previously circulated under the title "Corporate taxation, tax administration and financial development." I am grateful to David Agrawal, Katarzyna Bilicka, Marko Köthenbürger and Guttorm Schjelderup as well as to participants at the 2018 CESifo Area Conference on Public Sector Economics and the 74th annual congress of the International Institute of Public Finance for valuable discussions.

"...optimal policy requires simultaneous consideration of the design of the tax code and of the administrative structure created to enforce it."
– McLaren, (2003), p.v

1 Introduction

Tax rates and tax administrations' capacity are both critical elements to the design of optimal tax systems. This is especially the case in developing countries, where tax avoidance and evasion are even of greater concern than in developed countries. Despite the importance of tax administrations, they have received little attention in the literature. Accordingly, also the relationship between tax administrators and policymakers has not yet been discussed in the literature on tax systems. While tax administrations have traditionally been organized as an integral part of a ministry (Crandall, 2010), over the last decades, there has been a tendency, especially in developing countries, to establish so-called semi-autonomous revenue agencies thereby clearly demarcating competences of lawmakers and tax administrators (Junquera-Valera, 2019; Taliercio, 2004).

In this paper, I analyze the optimal tax system, consisting of the tax rate and the capacity of the tax administration, which determines subjects avoidance opportunities. The focus of my study is on carving out qualitative differences in the optimal tax system of developed and developing countries. The conventional wisdom is that tax rates and administrative capacities are complements, as suggested for example by Keen and Slemrod (2017) who write: "There are thus some grounds for a cautious presumption that (...) the best response to weaker administration is to set a lower tax rate (...)."

This presumption is, however, inconsistent with observed tax policies across countries. Administrative capacities are usually the lowest in developing countries, which is demonstrated by the wide variety of external supporters, such as the International Monetary Fund (IMF), the Organisation for Economic Co-operation and Development (OECD), the United Nations (UN) and the World Bank as well as bilateral donors such as Germany, Norway, the UK, and the USA that have been summoned by developing

countries to help build capacity in their tax administrations. However, Table 1 suggests, on average, for a variety of development indicators, a non-monotonic relationship between tax rates, both on corporate income and personal income, and development.

Table 1: Tax rates and development

Develop. ind.	q20	q40	q60	q80	q100	# countries
Panel A. Corporate income tax rates						
Economic	27.57	25.11	22.33	22.24	24.05	137
Financial	29,60	23.38	21.10	23.08	24.53	129
Human	29.15	26.67	23.60	19.31	24.37	120
Panel B. Top marginal personal income tax rates						
Economic	31.12	26.02	25.00	26.74	47.16	105
Financial	29.10	23.68	25.45	37.10	42.15	102
Human	28.87	29.50	24.90	27.80	43.97	98

Note: Samples are split into quintiles. All development indicators refer to the year 2017. Economic development is based on GDP per capita while financial development is measured as domestic credit provided to the private sector relative to GDP (both provided by the World Bank). Human development is based on the inequality-adjusted human development index provided by the United Nations. Corporate income tax rates are based on the year 2016 and retrieved from the Ernst & Young Worldwide Corporate Tax Guide and KPMG's corporate tax table, and personal income tax rates are based on the year 2017 and retrieved from KPMG's individual income tax rates table. Tax havens based on Hines (2005) are excluded from both panels, while Panel B additionally excludes countries that do not levy a personal income tax rate.

Based on the conventional wisdom that tax rates and administrative capacities are complements, it is therefore surprising to observe that the tax rates set by the countries with the lowest administrative capacity are actually among the highest.[1] This raises the question of whether the conventional wisdom of strategic complementary between

[1]The predominance of high tax rates in the least developed countries in Panel A does not originate from resource-rich countries' incentives to tax the rents from natural resource extraction. Although, in the financial development sample, 13 out of the top 30 countries in terms of income from natural resources (in % of GDP) belong to the first quintile (q20), all countries for which the E&Y World Corporate Tax Guide (2016) provides more details on industry specific rates (11 out of 13 countries) tax companies that are active in extractive industries at special rates.

policy instruments is as generally applicable as presumed.

In evaluating the optimality of the observed tax policies across countries, I first build on a two-country model featuring a small, non-haven country and a tax haven. A multinational enterprise has an operating affiliate (henceforth affiliate) in the non-haven country and a profit center in a tax haven and decides on the size of its investment in the affiliate as well as the amount of profit shifting to the tax haven. Using the terminology of Slemrod and Yitzhaki (2002), the multinational enterprise can therefore reduce its tax liability through reducing its capital investment (real response), and through profit relocation, which reduces the tax liability without necessarily altering the investment decision (avoidance). The non-haven country maximizes tax revenues by optimizing the tax system, that is by setting the level of the tax rate and the capacity of the tax administration, which affects the multinational enterprise's (marginal) costs of engaging in tax avoidance. Central to my analysis is the distinction between developed and developing countries, which I model by the presence of additional costs related to investing. A higher level of development, reduces these costs and thus increases incentives to invest.

Crucial to the evaluation of optimal tax policies is an aspect previously neglected in the tax systems literature: the tax administration's autonomy. In practice, tax administrators have varying degrees of freedom and are sometimes tightly accountable to lawmakers. If the tax administration is traditionally organized, that is an integral part of a ministry, usually the ministry of finance, the tax administration has little autonomy and is directly supervised by lawmakers (Crandall, 2010). In this case, the relationship between the tax rate as well as the tax administration's capacity and development is non-monotonic, where both developed and developing countries set higher tax rates and institute a higher administrative capacity than moderately developed countries. However, over the last decades, there has been a tendency, especially in developing countries, to increase the autonomy of their tax administration as a means to improve the effectiveness by establishing so-called semi-autonomous revenue agencies (Junquera-Varela, 2019; Taliercio, 2004). Typical powers of semi-autonomous revenue agencies encompass enforcement and penalties for acts of non-compliance, which the

tax administration can exercise without referral to another body. Such an institutional environment implies a two-stage game, where in the first stage one body, say the ministry of finance, determines the tax rate, whereas the tax administration determines its capacity in the second stage taking the tax rate as given. In this case, I show that the tax administration's capacity increases monotonically with the level of development, as suggested by the empirical literature, without affecting the non-monotonic relationship between the tax rate and development.

My results have important implications for taxation in developing countries. While there are good reasons to establish semi-autonomous revenue agencies, e.g. reducing the potential for corruption or productivity increases via management improvements (Crandall, 2010), my analysis suggests that they may actually be part of the cause of than the remedy for the low administrative capacity in developing countries. Instead, a significant improvement in developing countries' administrative capacities necessarily depends on following a more holistic approach, i.e. the simultaneous consideration of reforms to the tax system, as suggested by McLaren (2003) and more recently by the International Monetary Fund, the Organisation for Economic Co-operation and Development, the United Nations and the World Bank (IMF, OECD, UN and World Bank, 2016, p.24).

Finally, I shed light on the question whether fighting tax havens is beneficial for developed and developing countries. In order to do that, I first extend the model by a second non-haven country proving that the previously derived results are unaffected by the incorporation of tax competition. Based on this extension, I show that measures intended to reduce multinational enterprises' use of tax avoidance opportunities increase tax revenues in developed countries, because reduced profit shifting increases the tax base. Whether developing countries benefit from a fight against tax havens crucially depends on the institutional organization of the tax administration. If tax administrations have autonomy, fighting tax havens has an ambiguous effect on tax revenues as tax competition between the non-haven countries is affected in an ambiguous way. The result indicates that, although base erosion and profit shifting are a global concern, and even though initiatives, such as the OECD's Harmful Tax Practices Initiative, may have

been well-intentioned, the implications of base erosion and profit shifting concerns may be diverse for differently developed countries. If, instead, developing countries follow a more holistic approach, in which tax administrators and law makers tightly coordinate their actions, the fight against tax havens mitigates tax competition and thus also benefits them. Consequently, a holistic approach toward tax policies can eliminate the potential conflict of interest between developed and developing countries.

Despite the model's focus on corporate behavior, my analysis also applies to individual behavior and thus to other taxes, such as the personal income tax. Eventually, the main driver for my results is that real responses become more sensitive to taxation the lower the level of development, while a lower administrative capacity reduces the tax sensitivity. In the context of personal income taxes, individuals may choose to allocate working time in both the formal and informal sector similar to Sandmo (1981). Higher taxation in the formal sector should provoke a stronger reduction of individuals' time spend in the formal sector in developing countries as the relatively larger size of the informal sector offers more or comparably better outside options for individuals (real response channel). At the same time, as for example in Slemrod (2001), individuals may legally shelter some amount of income earned in the formal sector from taxation (avoidance channel). However, how much individuals are able to avoid depends on the capacity of the tax administration, which determines individuals' unavoidable compliance costs associated with taxpaying (Slemrod and Yitzhaki, 2002). A lower administrative capacity allows individuals to shelter a higher share of their tax payments, but may induce individuals to spend a higher fraction of their working time in the formal sector due to effectively lower tax payments.

2 Related literature

My study contributes to two strands of the literature. First, it relates to the literature on tax systems. Although tax reforms usually come with simulataneous changes in the tax rate and enforcement rules, the role of the tax administration in tax systems has

not yet received much attention.[2] Several contribution have highlighted the importance of limited administrative capacity (Gordon and Li, 2009; Dharmapala et al., 2011; Best et al., 2015; Kleven et al., 2016; Bachas and Soto, 2018), but usually focused on the implications for optimal taxation only. Other studies analyzed optimal taxation when governments can control tax avoidance (Mayshar, 1991; Slemrod, 1994; Slemrod and Yitzhaki, 2002; Keen and Slemrod, 2017).[3] However, whether tax systems differ or should differ between developed and developing countries has so far not been a topic in this literature strand.[4]

Second, my study relates to the literature, which deals with the question whether tax havens are beneficial or not for non-haven countries. While there is a continuing concern among policy makers that multinational enterprises' tax haven operations have a detrimental effect on non-haven countries' welfare because they erode the tax base (Slemrod and Wilson, 2009; Johannesen, 2012), there is also the alternative view that tax havens are beneficial because they allow governments to effectively tax discriminate between differently mobile firms (Hong and Smart, 2010), or reduce tax competition between non-haven countries (Johannesen, 2011).[5] My paper contributes to this literature by highlighting the importance of a country's level of development for whether tax havens are harmful.

3 The Model

3.1 The basic framework

I consider a one-period model of two small countries, a non-haven country, which levies a tax t, and a tax haven levying a tax t_0. Taxes are modeled as proportional taxes

[2]See Bird (1992, 2014) for a discussion of tax policy design and the relation to administration.

[3]See also Slemrod (2019) for a recent review of the literature on tax compliance and enforcement.

[4]See, for example, Emran and Stiglitz (2005) and Keen (2008) for contributions that highlight differences in tax practices among diversely developed countries. For more general discussions, see Besley and Persson (2013).

[5]For a critical summary of these views, see Dharmapala (2008).

on profits and are imposed at source.[6] The non-haven country determines its tax rate endogenously, whereas the tax rate of the tax haven, t_0, is exogenously given. There is one representative multinational enterprise (MNE) with an operating affiliate in the non-haven country and a profit center in the tax haven.[7] The affiliate produces a homogeneous good and sells it at the world market at a price normalized to one. The good is produced according to the technology $f(k)$, with positive but decreasing returns to capital, where k is capital investment. Moreover, for tractability $f'''(k) = 0$. Decreasing returns to scale in production imply the existence of a fixed factor, that is a firm-specific asset, e.g., a patent, that gives rise to positive pure profits.

I assume capital k to be the only production input. The MNE raises capital at an exogenous interest rate r, but investing capital is associated with additional costs for the MNE that depend on the host country's development. Specifically, investment costs are given by $\Omega = \frac{k}{\mu}$, where $\mu > 0$ is a parameter that is the higher the more developed the host country. Generally, the parameter μ captures costs indirectly related to the investment like the costs related to getting external finance, to enforce contracts or the lack of investor protection.

Additionally, the MNE may shift a fraction $\alpha \in [0, 1]$ of the affiliate's profits to its profit center located in the tax haven. Profit shifting is associated with a cost defined as $C = l\frac{(\alpha)^2}{2}\phi$, where ϕ is the affiliate's pre-shifting tax base as defined below in eq. (2). Concealment costs increase over-proportionally with the share of profits α the MNE shifts to the tax haven. Moreover, a higher capacity of the tax administration, l, makes it more costly for the MNE to shift profits abroad. Finally, the concealment costs also

[6] The source principle of taxation, where the profits of an affiliate are tax-exempt in the country of the parent firm, is followed by the majority of OECD countries, including the U.S. since 2018.

[7] In my analysis, I abstract from purely national firms. A justification for this assumption is that in developing countries the majority of firms are informal (Brockmeyer et al., 2015). Recent empirical evidence suggests that most informal firms will not formalize unless forced to do so (de Andrade et al., 2016). One explanation is that formal and informal firms operate in different markets and going formal at some point in time is usually not a widespread strategy among informal firms (La Porta and Shleifer, 2014). Moreover, the bulk of value added is produced by many small and a few large firms, a phenomenon called the "missing middle" (Dharmapala et al., 2011). Because small firms are usually exempt from taxation, the lion share of tax revenues comes from large, generally multinational, firms (Keen, 2012). This is the case even in developed countries where a larger fraction of value added is produced by medium-sized firms.

imply that if the MNE does not shift profits ($\alpha = 0$) or the tax administration does not control profit shifting, ($l = 0$), (marginal) costs of profit shifting are zero.

3.2 The multinational enterprise

Economic profits of the MNE in country i are given by the income from selling output less the user cost of capital

$$\pi = f(k) - rk - \frac{k}{\mu}. \tag{1}$$

For simplicity, I assume that capital costs are fully tax-deductible, but costs indirectly related to investing in the country are not tax-deductible. Hence, taxable profits are given by

$$\pi^t = (1-\alpha)[f(k) - rk] \equiv (1-\alpha)\phi. \tag{2}$$

Because the MNE can shift a fraction of the affiliate's profits to the tax haven, total after-tax profits of the MNE read

$$\Pi = \psi\phi - \frac{k}{\mu}. \tag{3}$$

where $\psi = \left[(1-t)(1-\alpha) + (1-t_0)\alpha - l\frac{(\alpha)^2}{2}\right]$. The MNE maximizes its profits by choosing its investment level and the share of profits shifted to the tax haven. The optimal level of capital investment is given by

$$\psi[f'(k) - r] - \frac{1}{\mu} = 0. \tag{4}$$

From the first-order condition, I can derive the effects of the tax rate, the administrative capacity and the country's level of development on the optimal capital investment. Totally differentiating equation (4) yields

$$\frac{\partial k}{\partial t} = \frac{(1-\alpha)[f'(k) - r]}{\psi f''(k)} < 0, \quad \frac{\partial k}{\partial l} = \frac{\alpha^2[f'(k) - r]}{2\psi f''(k)} < 0,$$

$$\frac{\partial k}{\partial \mu} = -\frac{1}{\mu^2 \psi f''(k)} > 0. \tag{5}$$

A higher tax rate t reduces the incentives to investment in the non-haven country. However, the tax sensitivity of capital depends on the MNE's ability to shift profits. The larger the amount of profits the MNE can shift to the tax haven, the lower the tax sensitivity of capital. This is in line with empirical evidence by, for example, Overesch (2009) and Büttner et al. (2018) who find that profit shifting lowers the tax sensitivity of real investment. Second, a higher administrative capacity increases the costs of shifting profits, which leads to a lower capital investment because a higher fraction of the return on the investment is taxed at the higher tax rate of the non-haven country instead of the tax haven's. Third, a higher level of development reduces the user cost of capital and thus increases capital investment in the non-haven country.

The optimal share of profits shifted to the tax haven is determined by

$$\alpha = \frac{t - t_0}{l}, \tag{6}$$

which states that the MNE shifts profits from its affiliate to its profit center until the marginal tax savings are equal to the marginal concealment costs. From equation (6), I can derive the effects of changes in the tax rates and the tax administrations' effectiveness on α, which are given by

$$\frac{\partial \alpha}{\partial t} = \frac{1}{l} > 0, \quad \frac{\partial \alpha}{\partial l} = -\frac{t - t_0}{(l)^2} = -\frac{\alpha}{l} < 0. \tag{7}$$

A higher tax rate in the non-haven country increases the incentives of the MNE to shift profits to the tax haven. Instead, a higher administrative capacity reduces profit shifting because the associated costs go up.

3.3 Government

Turning to the government, I assume its objective to be tax revenue maximization, an assumption frequently made in the international tax literature. However, in the context of my analysis, two additional reasons can be brought forward to justify this assumption. First, tax revenue considerations play an important role in developing countries,

which usually raise a large fraction of their revenues from the taxation of MNEs. Also for developed countries, tax revenue considerations are relevant not only because of severe revenue shortfalls due to profit shifting, but also due to equality-of-treatment concerns. Together these concerns exert strong political as well as practical pressures on governments to increase revenues from this source.[8] Second, tax revenue maximization abstracts from other channels that may affect tax policies differently in developed and developing countries. For example, the extent to which policies are affected by labor effects related to attracting foreign direct investment or the opportunity of corruption, lobbying or favorable policies toward a ruling elite may culminate in welfare functions differing between developed and developing countries. I show that the non-monotonic relationship displayed in Table 1 can be explained even in the absence of such differences.

Tax revenues are given by

$$T = t(1-\alpha)\phi - \kappa(l), \tag{8}$$

where $\kappa(l)$ is the cost of increasing the capacity of the tax administration and is assumed to be convex, i.e. $\kappa'(l), \kappa''(l) > 0$. I further assume that $\kappa'(0) = 0$, which ensures at least some extent of profit shifting control.

Differentiating tax revenues with respect to t and l implicitly determines the optimal tax system

$$\frac{\partial T}{\partial t} = (1-\alpha)\phi + \frac{t(1-\alpha)^2[f'(k)-r]^2}{\psi f''(k)} - \frac{t\phi}{l} = 0, \tag{9}$$

$$\frac{\partial T}{\partial l} = t\left[\frac{(1-\alpha)\alpha^2[f'(k)-r]^2}{2\psi f''(k)} + \frac{\phi\alpha}{l}\right] - \kappa'(l) = 0. \tag{10}$$

The first-order condition for the optimal tax rate (9) states that raising t increases tax revenues due to increased tax payments by the affiliate (first term). However, a higher tax rate also reduces tax revenues because of lower investments by the affiliate (second term) and more profit shifting to the tax haven (third term). Evaluating equation (9) at $t = 0$ shows that the second and third terms vanish and the derivation is positive

[8] See, for example, the motivation for the OECD's base erosion and profit shifting (BEPS) initiative (OECD, 2013, Chapters 1 and 2).

at $t = 0$, which results in $t > 0$ in equilibrium.

The first-order condition for the optimal capacity of the tax administration (10) states that increasing l reduces tax revenues because it decreases the affiliate's capital investment due to the lower net return on capital (first term). However, it also increases tax revenues due to the reduction in profit shifting (second term). Evaluating condition (10) at $l = 0$ shows that the first term vanishes because α approaches 1. Since the costs of increasing the capacity of the tax administration are negligible for $l = 0$, i.e. $\kappa'(0) = 0$, the first-order condition is positive at $l = 0$, which results in $l > 0$ in equilibrium.

The previous analysis shows that the non-haven country has an incentive to limit the amount of profit shifting by the MNE in order to tax the income generated through production. However, the extent of profit shifting control and taxation may vary with the country's level of development. I analyze this question in the next section.

4 Optimal tax policies and development

In this section, I analyze how a change in the non-haven country's development affects the tax rate and the capacity of the tax administration. The purpose of this exercise is to evaluate whether optimal tax systems differ between developed and developing countries. Hence, the main interest lies in determining the signs of $\frac{dt}{d\mu}$ and $\frac{dl}{d\mu}$. Totally differentiating the first-order conditions for t and l leads to

$$\begin{bmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{bmatrix} \begin{bmatrix} dt \\ dl \end{bmatrix} = \begin{bmatrix} -\theta_5 \\ -\theta_6 \end{bmatrix} d\mu, \qquad \text{where} \tag{11}$$

$$\theta_1 = \frac{\partial^2 T}{\partial t^2}, \quad \theta_2 = \frac{\partial^2 T}{\partial t \partial l} = \frac{\partial^2 T}{\partial l \partial t} = \theta_3, \quad \theta_4 = \frac{\partial^2 T}{\partial l^2}, \quad \theta_5 = \frac{\partial^2 T}{\partial t \partial \mu}, \quad \theta_6 = \frac{\partial^2 T}{\partial l \partial \mu}, \tag{12}$$

where Appendix A.1 provides details on the derivation of the θ-terms. Applying Cramer's rule to the equation system given in (11) yields

$$\frac{dt}{d\mu} = \frac{\theta_2\theta_6 - \theta_4\theta_5}{|A|}, \qquad \frac{dl}{d\mu} = \frac{\theta_2\theta_5 - \theta_1\theta_6}{|A|}, \tag{13}$$

where $|A| = \theta_1\theta_4 - (\theta_2)^2 > 0$ to obtain a maximum. In general, the effect of an improvement in the development level, that is an increase in μ, on the tax rate and the tax administration's capacity is ambiguous because the sign of θ_2, that is the relation between policy instruments, is undetermined.

To draw inferences about the effects given in (13), I make the following distinction between developed and developing countries: the fact that an improvement of development, that is an increases in μ, implies a higher capital investment and is therefore associated with a lower marginal net return on investment $f'(k) - r$. Hence, in developing countries the marginal net return on investment $f'(k) - r$ is high, whereas it is low in developed countries.

Using this differentiation, Appendix A.2 shows that, starting from a high level of μ, the effects on the tax rate and the tax administration's capacity are given by

$$\frac{dt}{d\mu} > 0, \qquad \frac{dl}{d\mu} > 0. \tag{14}$$

The comparative static results in (14) state that, conditional on a high level of μ, an improvement of development increases both the tax rate and the tax administration's capacity. The intuition for these results originates from the fact that tax base effects are much more relevant than base sensitivity effects in developed countries. This implies that the tax rate and the effectiveness of the tax administration are complements ($\theta_2 > 0$), i.e. the higher the level of the tax rate, the stronger the incentive to increase the tax administration's capacity, and vice versa. Thus, besides the direct incentive to raise the level of both policy instruments, which is cause by the higher investment due to the improvement of development (rise in μ), this incentive is reinforced indirectly because of the strategic complementarity of the two policy instruments.

Next, I analyze how the tax rate and the tax administration's capacity are affected if the initial level of development is low (low μ). Appendix A.2 shows that the effects are given by

$$\frac{dt}{d\mu} < 0, \qquad \frac{dl}{d\mu} < 0. \tag{15}$$

The results state that both the tax rate and the tax administration's capacity decline if the level of development rises. These results originate from the fact that base sensitivity effects are much more relevant than tax base effects in developing countries. Contrary to developed countries, this implies that the tax rate and the tax administration's capacity are substitutes, i.e. the higher the level of the tax rate the lower the capacity of the tax administration and vice versa.

The reason for the negative relationship between policy instruments and development is that the indirect effect dominates the direct effect. Specifically, an increase in μ directly increases the incentive to raise the level of either policy instrument due to higher investments by the MNE. However, this implies a drastic outflow of capital due to the high sensitivity of capital (cf. eq. (5)). Because policy instruments are substitutes ($\theta_2 < 0$), the consequence of the drastic reduction in the tax base is a strong incentive to lower the level of both policy instruments, which overcompensates the initial incentive to raise them.

Combining the two results, the analysis indicates a non-monotonic relationship between policy instruments and development. At low levels of development tax rates and tax administrations' capacity decrease with development, whereas at high levels of development further improvements lead to a higher tax rate and a higher capacity of tax administrations. I summarize in:

Proposition 1 *The relationship between development and tax rates as well as the tax administrations' capacity is non-monotonic.*

Proposition 1 challenges the conventional wisdom that tax rates and administrative capacities are complements. Indeed, the non-monotonic relationship between policy

instruments and development only arises because the tax rate and the tax administration's capacity are substitutes in developing countries. My results show the non-monotonic relationship displayed in Table 1 is not coincidental, but emerges from countries' optimization of their tax system given their stage of development.

However, in contrast to the empirical literature (Fuest et al., 2011; Crivelli et al., 2016; Johannesen et al., 2017), Proposition 1 states that the optimal tax system in developing countries should feature a higher capacity of the tax administrations than comparably more developed countries. Against this background, the question arises why, in practice, administrative capacities remain so weak in developing countries. If we look at the first-order condition for the optimal capacity of the tax administration in isolation, equation (10) indeed shows that less developed countries have a lower incentive to improve their tax administration's capacity ($\theta_6 > 0$). The reason is that a lower level of development implies a lower level of capital investment, and thus a lower pre-shifting base ϕ, but a higher capital sensitivity. This means that the magnitude of the negative first effect in (10) is relatively larger in less developed countries, whereas the positive second effect is smaller, which culminates in a lower capacity of the tax administration. Hence, this suggests that the administrative capacity is determined as if it was an isolated choice, i.e. as if the indirect effect via the change in the tax rate, which is the reason for the non-monotonic relationship between the tax administration's capacity and development, did not matter.

What remains to be answered is why, in practice, this is the case. One obvious reason is that tax policies are determined in a two-stage process, where in the first stage the tax rate is determined by one agency, say the Ministry of Finance, taking into account the reaction of the tax administration, which, in the second stage, determines its capacity, taking the decisions made in the first stage as given.

Indeed, over the last decades, there has been a global tendency to create semi-autonomous revenue agencies (SARA) as a means to improve the administration of tax systems (Junquera-Varela, 2019; Taliercio, 2004). While day-to-day operations are delegated to SARA, which are directly accountable for their implementation, the formulation of tax policy remains with the Ministry of Finance (Terkper, 2008). Typical

powers of SARA encompass enforcement and penalties for acts of non-compliance, which they can exercise without referral to another body. By 2012, 31 out of 52 countries surveyed by the OECD have established a SARA responsible for tax administration operations (OECD, 2013). In Latin America and Africa, establishing SARA is the predominant pattern of tax administration reform (Crandall, 2010), a feature related to developing countries summoning a wide variety of external supporters, such as the IMF, the OECD, the UN and the World Bank as well as bilateral donors such as Germany, Norway, the UK, and the USA, to help build capacity in their tax administrations.

If the tax administration's capacity is determined in the second stage the effect of a change in development on the optimal capacity is given by

$$\frac{dl}{d\mu} = -\frac{\theta_6}{\theta_4} > 0. \tag{16}$$

Importantly, because the capacity of the tax administration is not taken as given when the tax rate is determined, the non-monotonic relationship between tax rates and development remains qualitatively unaffected. I summarize in:

Proposition 2 *If the tax administration's capacity is determined after the tax rate (two-stage process), the non-monotonic relationship between the tax rate and development remains unaffected, but the tax administration's capacity increases monotonically with development.*

Despite great differences in administrative capacities between developed and developing countries, many external supporters are rooted in OECD countries offering advice that usually fits with their own values and priorities and not with those of developing countries (Moore et al., 2015). Crandall (2010) points out that, among other things, it is critical to understand the benefits and downsides of SARA as well as whether the conditions for success and sustainability are present before making any decision to proceed with their establishment.

Proposition 2 highlights that the seemingly suboptimal combination of comparably high tax rates and low administrative capacities in developing countries is actually

the outcome of tax system optimization when tax administrations have autonomy. Thus, although there are good reasons to establish SARA in order to improve tax administrations' capacity, e.g. reducing the potential for corruption or productivity increases via management improvements (Crandall, 2010), Proposition 2 suggests that SARA may actually be part of the cause of than the remedy for the weak administrative structures in developing countries.

While the existing empirical literature largely finds a positive correlation between the establishment of SARAs and increases in revenue collection, Dom (2018) argues these estimates are biased as they do not account for differences in pre-reform trends in revenue collection. Indeed, controlling for pre-trends, Dom (2018) concludes there is no systematic relationship between the presence of a SARA and total tax revenue in sub-Saharan Africa. Thus, if external supporters provide financial assistance that is intended to increase administrative capacities in developing countries, it may be the case that this assistance does not fall on fertile ground due to lacking incentives.[9]

Instead, Proposition 1 emphasizes that a necessary condition for implementing a higher administrative capacity in developing countries, is that tax administrators and lawmakers tightly coordinate their actions as suggested by McLaren (2003). Against the background that developing countries already request external supporters to assist them in improving their administrative capacity, such an engagement could offer a unique opportunity to coordinate all directly tax-relevant agencies, like the finance department and the tax administration. In fact, such a holistic approach has recently been put forward jointly by the IMF, the OECD, the United Nations and the World Bank as a cornerstone to commit to formulated medium-term revenue strategies (see IMF, OECD, UN and World Bank (2016), p.24 as well as Recommendation 2d).

[9]Such a problem has already been documented for the case of foreign aid when money is not spent on the purpose for which it was intended (e.g. Andersen et al., 2020).

5 Extensions

In this section, I show that incorporating tax competition does not change the results derived in the previous section. Using the tax competition framework, I will then analyze the implications of fighting tax havens for developed and developing countries.

5.1 Tax competition

I extend the model described in section 3 by introducing a second non-haven country, in which the MNE has an affiliate that produces output. While the basic framework remains the same in this section, there is one main difference to the previous analysis with only one affiliate. In particular, affiliates have to raise capital at an exogenous interest rate r from national credit markets. However, the total amount of capital the MNE can allocate between affiliates is insufficient to ensure that the marginal return on investment in each affiliate equates the marginal costs in the optimum. By how much affiliates' marginal return on investment exceeds the marginal cost, i.e. how severe affiliates' finance constraint is, depends on countries' level of development.

While the empirical literature has shown that MNEs can use their internal capital market to mitigate their affiliates' finance constraint (Manova et al., 2011), they still play an important role for them. Desai et al. (2004), for example, find that the use of internal resources can only partially compensate the lack of external finance in countries with a weak financial development. Antràs et al. (2009) show that, affiliates' sales decrease with the magnitude of financial underdevelopment despite the higher reliance on parental finance. More generally, Egger et al. (2014) show that investments of different affiliates within the same multinational group are interrelated in the presence of an internal capital market, which is indicative of scarce internal resources. The MNE's resource constraint is thus given by $k_i + k_j = \bar{k}$, where $\bar{k}$ is the MNE's total amount of capital and determined by the development of the countries.[10]

[10]This modeling is a reduced form of a standard moral hazard problem in which an agent (or a manager) chooses an unobservable effort level, which also influences the success probability of an investment, and external lenders restrict their funds to incentivize the manager. A moral-hazard based

Because the MNE can shift a fraction of the profits of affiliate i, $i = \{a, b\}$, to the tax haven, total after-tax profits of the MNE read

$$\Pi = \psi_i \phi_i + \psi_j \phi_j. \tag{17}$$

where $\psi_i = \left[(1-t_i)(1-\alpha_i) + (1-t_0)\alpha_i - l_i \frac{(\alpha_i)^2}{2}\right]$. Since the MNE's total capital is fixed to $\bar{k}$, capital is allocated to each affiliate such that in equilibrium the net returns are equalized across the two affiliates, that is

$$\psi_i[f'(k_i) - r] = \psi_j[f'(k_j) - r] > 0. \tag{18}$$

From the first-order condition (18), I can derive the effects of tax rates and administrative capacities on optimal capital investment in countries i and j. Appendix B.1 shows that totally differentiating equation (18) yields

$$\begin{aligned}
\frac{\partial k_i}{\partial t_i} &= \frac{(1-\alpha_i)[f'(k_i) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)} < 0, \qquad \frac{\partial k_i}{\partial t_j} = -\frac{(1-\alpha_j)[f'(k_j) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)} > 0, \\
\frac{\partial k_i}{\partial l_i} &= \frac{\frac{(\alpha_i)^2}{2}[f'(k_i) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)} < 0, \qquad \frac{\partial k_i}{\partial l_j} = -\frac{\frac{(\alpha_j)^2}{2}[f'(k_j) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)} > 0.
\end{aligned} \tag{19}$$

The interpretation of equation (19) is the same as for equation (5) with the only difference that raising the tax rate or the tax administration's capacity leads to an outflow of capital to the other non-haven country.

The optimal share of profits shifted from affiliate i to the tax haven does not change and is thus still given by (6). However, it is worth mentioning that the foreign tax rate t_j and the capacity of the foreign tax administration l_j do not affect the MNE's profit shifting incentives related to affiliate i.

While tax revenues are not affected qualitatively, the first-order conditions for t_i and l_i slightly change due to the presence of tax competition. Differentiating tax revenues

relationship between the capital market and the financial development can be found in Holmstrom and Tirole (1997).

with respect to t_i and l_i yields

$$\frac{\partial T_i}{\partial t_i} = (1-\alpha_i)\phi_i + t_i \frac{(1-\alpha_i)^2[f'(k_i)-r]^2}{\psi_i f''(k_i)+\psi_j f''(k_j)} - t_i \frac{\phi_i}{l_i} = 0, \tag{20}$$

$$\frac{\partial T_i}{\partial l_i} = t_i \left[\frac{(1-\alpha_i)[f'(k_i)-r]^2 \frac{(\alpha_i)^2}{2}}{\psi_i f''(k_i)+\psi_j f''(k_j)} + \phi_i \frac{\alpha_i}{l_i} \right] - \kappa'(l_i) = 0. \tag{21}$$

Using the simplifying assumption that capital mobility is frictionless, which implies a symmetric outcome despite initial asymmetries in countries' level of development, Appendix B.2 shows that totally differentiating the first-order conditions yields

$$\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_2 & \theta_6 & \theta_7 & \theta_8 \\ \theta_3 & \theta_4 & \theta_1 & \theta_2 \\ \theta_7 & \theta_8 & \theta_2 & \theta_6 \end{bmatrix} \begin{bmatrix} dt_i \\ dl_i \\ dt_j \\ dl_j \end{bmatrix} = \begin{bmatrix} -\theta_{17} \\ -\theta_{18} \\ -\theta_{17} \\ -\theta_{18} \end{bmatrix} d\bar{k}, \quad \text{where} \tag{22}$$

$$\theta_1 = \frac{\partial^2 T_i}{\partial (t_i)^2}, \quad \theta_2 = \frac{\partial^2 T_i}{\partial t_i \partial l_i}, \quad \theta_3 = \frac{\partial^2 T_i}{\partial t_i \partial t_j} \quad \theta_4 = \frac{\partial^2 T_i}{\partial t_i \partial l_j}, \quad \theta_6 = \frac{\partial^2 T_i}{\partial (l_i)^2},$$
$$\theta_7 = \frac{\partial^2 T_i}{\partial l_i \partial t_j}, \quad \theta_8 = \frac{\partial^2 T_i}{\partial l_i \partial l_j}, \quad \theta_{17} = \frac{\partial^2 T_i}{\partial t_i \partial \bar{k}}, \quad \theta_{18} = \frac{\partial^2 T_i}{\partial l_i \partial \bar{k}}.$$

When analyzing how a change in the level of development affects tax policies, I draw the same distinction between developed and developing countries as in section 4: the higher the level of development ($\bar{k}$), the higher the investment level and thus the lower the net marginal return on investment. Using this differentiation, Appendix B.3 shows that, starting from a high level of $\bar{k}$, the effects on the tax rates and tax administrations' capacity are given by

$$\frac{dt_i}{d\bar{k}} > 0, \qquad \frac{dl_i}{d\bar{k}} > 0. \tag{23}$$

In contrast, starting from a low level of $\bar{k}$, the effects on the tax rates and tax admin-

istrations' capacity read[11]

$$\frac{dt_i}{d\bar{k}} < 0, \qquad \frac{dl_i}{d\bar{k}} < 0. \tag{24}$$

The intuition for the results is the same as in section 4. The change in the sign of the derivations arises because policy instruments are complements in developed countries, while they are substitutes in developing countries.

If instead tax policies are determined in a two-stage process, Appendix B.3.2 shows that the effect of a change in the level of development on tax administrations' capacity is given by

$$\frac{dl_i}{d\bar{k}} > 0. \tag{25}$$

Hence, tax administrations' capacity increases monotonically with development. Taken together, the analysis shows that the presence of tax competition does not qualitatively alter the conclusions summarized in Propositions 1 and 2.

5.2 The implications of fighting tax havens

Based on the insight that the incorporation of tax competition does not change the previously derived results, I will use this extended framework to analyze the implications of fighting tax havens. Propositions 1 and 2 have shown that critical elements that distinguish optimal tax systems in developed and developing countries is the relationship between tax rates and administrative capacities as well as the degree of tax administrations' autonomy. Based on these aspects, the fight against tax havens may have different implications for countries in different stages of their development.

In the following, I therefore analyze how countries' tax revenues will be affected by an intensified fight against tax havens. I interpret actions, which aim to reduce the preferential tax treatment offered by tax havens as a rise in the exogenously given tax

[11]For reasons of space, I have omitted the proof for developing countries because it is lengthy and tedious. However, the complete derivation is available on my homepage.

rate t_0. Eventually, such practices imply weaker incentives for the MNE to shift profits to the tax haven. The total effect of an increase in t_0 on tax revenues in country i is given by

$$\frac{dT_i}{dt_0} = \frac{\partial T_i}{\partial t_0} + \frac{\partial T_i}{\partial t_j}\frac{\partial t_j}{\partial t_0} + \frac{\partial T_i}{\partial l_j}\frac{\partial l_j}{\partial t_0}, \tag{26}$$

where the first term in equation (26) is the direct effect of a change in the tax haven's tax rate on country i's tax revenues, while the last two terms capture the implications of a change in t_0 for tax competition.

Differentiating equation (8) with respect to t_0, and taking into account that the MNE allocates capital equally between the two affiliate given that capital mobility is frictionless, this implies that $\frac{\partial k_i}{\partial t_0} = 0$. Thus, the direct effect of an increase in t_0 reads

$$\frac{\partial T_i}{\partial t_0} = \frac{t\phi}{l} > 0. \tag{27}$$

Equation (27) shows that the direct effect is unambiguously positive. The reason is that a higher tax rate in the tax haven reduces profit shifting by the MNE and thus leads to a larger tax base in country i.

Next, I determine how the fight against tax havens affects tax competition between non-haven countries. To do that, I start by deriving the fiscal externalities country j exerts on country i. Differentiating equation (8) with respect to t_j and l_j respectively delivers

$$\frac{\partial T_i}{\partial t_j} = t_i(1-\alpha_i)[f'(k_i) - r]\frac{\partial k_i}{\partial t_j} > 0, \tag{28}$$

$$\frac{\partial T_i}{\partial l_j} = t_i(1-\alpha_i)[f'(k_i) - r]\frac{\partial k_i}{\partial l_j} > 0. \tag{29}$$

The two externalities are positive and state that an increase both in the tax rate and in the tax administration's capacity in country j raise country i's tax revenues. The reason is that a higher tax rate or a higher capacity of the tax administration in country j reduces the net-of-tax return on investment in country j and thus induces the MNE to reallocate capital from country j to country i. The higher investment in country i increases the tax base and in turn tax revenues.

Plugging (27), (28) and (29) into equation (26), using (19) and that capital mobility is frictionless, the total effect on tax revenues is given by

$$\frac{dT_i}{dt_0} = \frac{t\phi}{l} - \frac{t(1-\alpha)[f'(k)-r]^2}{2\psi f''(k)} \left[(1-\alpha)\frac{\partial t_j}{\partial t_0} + \frac{\alpha^2}{2}\frac{\partial l_j}{\partial t_0} \right]. \tag{30}$$

Whether the direct effect or the indirect effects determine the sign of equation (30) depends on the level of development. If the level of $\bar{k}$ is high, the pre-shifting tax base ϕ is large, whereas the rate of return $[f'(k) - r]$ is low. Hence, for a sufficiently high development, the first term in equation (30), that is the direct effect, will dominate. Since this term is positive, policies that restrict the use of tax havens result in higher tax revenues for developed countries. I summarize in:

Proposition 3 *Fighting tax havens increases tax revenues in developed countries.*

What remains to be determined, is how a change in t_0 affects governments' tax revenues in developing countries. As the level of $\bar{k}$ in developing countries is low, and the marginal return on investment is therefore high, the second term in equation (30), i.e. the effect of a change in t_0 on tax competition, will determine the sign.

Based on the insight of Proposition 2 that tax policies are often determined in a two stage process, Appendix B.4 shows that an increase in t_0 when $\bar{k}$ is sufficiently low has the following effect on the capacity of the tax administration:

$$\frac{dl_j}{dt_0} < 0. \tag{31}$$

Fighting tax havens lowers the tax administrations' capacity in developing countries. The main reason is that less profit shifting, i.e. a lower level of α leads to a higher tax elasticity of capital (cf. (19)), which tax administrations counter by lowering their capacity in order to moderate the increase in tax competition.

While it is not possible to draw final conclusions about how a change in t_0 affects the tax rate, there is a tendency for developing countries to raise their tax rate following an increase in t_0 if $f''(k)$ is not too high.[12] The main reason for the tendency toward higher

[12]For reasons of space, I have omitted this proof, but it is available on the my homepage.

tax rates lies in the relationship between the two policy instruments. Specifically, the increase in t_0 provides a direct incentive to reduce the tax rate because a lower share of affiliate i's profits shifted to the tax haven increases the tax-sensitivity of capital (cf. (19)). However, because an increase in t_0 also reduces tax administrations' capacity (cf. (31)), this provides an incentive to increase the tax rate because the two policy instruments are substitutes. This latter effect dominates culminating in a tendency to increase the tax rate following a rise in t_0.

Overall, fighting tax havens has an ambiguous effect on tax competition and therefore on tax revenues in developing countries as tax rates and tax administrations' capacity are affected in opposite ways. I summarize in:

Proposition 4 *If tax administrations' capacity is determined after the tax rate (two-stage process), fighting tax havens has an ambiguous effect on developing countries' tax revenues.*

Propositions 3 and 4 can be related to various countries' and supranational organizations' initiatives against tax havens. One of the most prominent initiative is the OECD's Harmful Tax Practices Initiative the aim of which is to increase the costs for a country of offering tax sheltering opportunities (OECD, 1998). Since then, many of the jurisdictions deemed as tax havens have agreed to make commitments to implement the OECD's standards of transparency and exchange of information. The OECD's intention behind the shutting-down of countries' tax haven activities is to curtail the depletion of OECD countries' tax bases by discouraging foreign investors from using preferential tax regimes.

However, the results of Propositions 3 and 4 highlight that whether countries benefit from such a course of action to a great extent depends on their level of development. According to Proposition 4, MNEs' reduced possibilities of shifting their profits to tax havens decreases the erosion of developed countries' tax bases and thus result in higher tax revenues. Based on this result, it is therefore not surprising that the OECD's initiative has been driven by the interests of developed countries, whereas the interests of non-member countries, especially those of developing countries, are usually not being

addressed (Oguttu, 2016). Proposition 4 highlights that the outcome of a fiercer fight against tax havens may have different tax revenue effects for developing countries. In the light of developing countries' need for resources to foster economic growth, the potential of a shortfall in tax revenues following the fight against tax havens may make them even more dependent on outside financial assistance like development aid.

Finally, I analyze the implications of fighting tax havens when a more holistic approach is followed in developing countries, as recently suggested by the main supranational organizations (see, IMF, OECD, UN and World Bank, 2016). If tax policies are determined simultaneously, the second term in equation (30) is positive under the sufficient and very mild condition that the equilibrium tax rate is not too high ($t < 0.8$).[13] This will be the case if $f''(k)$ is not too high, that is if tax competition is sufficiently strong. As before the main reason originates from tax policies being substitutes in developing countries. The only difference compared to the previous case is that under a holistic approach, there is also an incentive to increase the tax administration's capacity, because indirect effects via the tax rate are taken into consideration. Overall, the fight against tax havens implies a reduction in tax competition and hence a rise in tax revenues if $f''(k)$ is sufficiently small. I summarize in:

Proposition 5 *Fighting tax havens increases tax revenues in developing countries if tax administrators and lawmakers tightly coordinate their actions and tax competition is sufficiently fierce.*

Although initiatives against tax havens may be primarily driven by the interests of developed countries, Proposition 5 illustrates that the potential conflict of interest between developed and developing countries can be resolved if developing countries followed a more holistic approach. Consequently, Proposition 5 also indicates that tight coordination of all directly tax-relevant agencies, like the finance department and the tax administration, in developing countries may feature a double dividend by not only laying the foundation for sustainable improvements in administrative capacity, but also by aligning heterogeneous countries' preferences for international policy cooperation.

[13] Again, I have omitted the proof because of its length, but the complete derivations are available on my homepage.

6 Conclusions

In this paper, I analyze the optimal tax system consisting of two elements, the level of the tax rate and the tax administration's capacity, and shed light on differences in tax systems across countries of varying levels of development. First, I challenge the conventional wisdom that the tax rate and the tax administrative capacity are complements by showing that their relationship depends on the level of development. While they are complements in developed countries, they are substitutes in developing countries. Second, a crucial aspect previously neglected in the tax systems literature is the tax administration's autonomy. When tax administrations are traditionally organized, that is directly supervised by lawmakers and therefore with little autonomy, the relationship between each policy instrument and development is non-monotonic, with both developed and developing countries setting higher tax rates and have higher administrative capacities than moderately developed countries. In contrast, when tax administrations have autonomy, as suggested by the developments over the last decades, the tax administration's capacity increases monotonically with the level of development, as suggested by the empirical literature, without affecting the non-monotonic relationship between tax rates and development.

My findings have two important policy implications. First, when tax administrations have autonomy, financial assistance provided by external supporters intended to improve tax administrations' capacity in developing countries may not fall on on fertile grounds due to lacking incentives. Second, in order to significantly improve developing countries administrative capacity, a holistic approach, that is tight coordination between tax administrators' and lawmakers' actions, is necessary as recently suggested by the IMF, the OECD, the United Nations and the World Bank (IMF, OECD, UN and World Bank, 2016).

Furthermore, I analyze, in a model capturing tax competition, whether fighting tax havens is beneficial for non-haven countries from a tax revenue perspective. While tax revenues in developed countries increase due to reduced profit shifting opportunities of multinational enterprises, such a practice have an ambiguous effect on tax revenues

in developing countries if tax administrations are autonomous. However, if developing countries follow a more holistic approach, the fight against tax havens will also be beneficial for them.

References

Andersen J.J., N. Johannesen and B. Rijkers, 2020. Elite capture of foreign aid: Evidence from offshore bank accounts. Policy Research Working Paper 9150.

de Andrade G.H., M. Bruhn, D. McKenzie, 2016. A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. The World Bank Economic Review 30, 24-54.

Bachas P. and M. Soto, 2018. Not(ch) your average tax system: Corporate taxation under weak enforcement. Policy Research Working Paper 8524.

Besley T. and T. Persson, 2013. Taxation and development. Handbook of Public Economics 5, 51-110.

Best M., A. Brockmeyer, H. Kleven, J. Spinnewijn and M. Waseem, 2015. Production versus revenue efficiency with limited tax capacity: Theory and evidence from Pakistan. Journal of Political Economy 123, 1311-1355.

Bird R.M., 1992. Improving tax administration in developing countries. Journal of Tax Administration 1, 23-45.

Bird R.M., 2014. Administrative dimensions of tax reform. Annals of Economics and Finance 15-2, 269-304.

Brockmeyer A., S. Smith, M. Hernandez and S. Kettle, 2019. Casting a wider tax net: Experimental evidence from Costa Rica. American Economic Journal: Economic Policy 11, 55-87.

Büttner T., M. Overesch, and G. Wamser, 2018. Anti profit-shifting rules and foreign direct investment. International Tax and Public Finance 25, 553-580.

Crivelli E., R. De Mooij and M. Keen, 2016. Base erosion, profit shifting and developing countries. FinanzArchiv/Public Finance Analysis 72, 268-301.

Desai M.A., C.F. Foley, J.R. Hines, 2004. A multinational perspective on capital structure choice and internal capital markets. Journal of Finance 59, 2451-2487.

Dharmapala D., 2008. What problems and opportunities are created by tax havens. Oxford Review of Economic Policy 24, 661-679.

Dharmapala D., J. Slemrod and J.D. Wilson, 2011. Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs. Journal of Public Economics 95, 1036-1047.

Egger P., C. Keuschnigg, V. Merlo and G. Wamser, 2014. Corporate taxes and internal borrowing within multinational firms. American Economic Journal: Economic Policy 6, 54-93.

Emran M.S. and J.E. Stiglitz, 2005. On selective indirect tax reform in developing countries. Journal of Public Economics 89, 599-623.

Fuest C., S. Hebous and N. Riedel, 2011. International debt shifting and multinational firms in developing economies. Economics Letters 13, 135-138.

Gordon R. and W. Li, 2009. Tax structures in developing countries: Many puzzles and a possible explanation. Journal of Public Economics 93, 855-866.

Hines J.R., 2005. Do tax havens flourish? in: Poterba J.M. (ed.), Tax policy and the economy 19, Cambridge: MIT Press.

Holmstrom B. and J. Tirole, 1997. Financial intermediation, loanable funds and the real sector. Quarterly Journal of Economics 112, 663-691.

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

IMF, OECD, UN and World Bank, 2011. Supporting the development of more effective tax systems. Report to the G-20 Development Working Group. Washington DC/Paris/New York.

IMF, OECD, UN and World Bank, 2016. Enhancing the effectiveness of external support in building tax capacity in developing countries. Prepared for Submission to G-20 Finance Ministers. Washington DC/Paris/New York.

Johannesen N., 2010. Imperfect tax competition for profits, asymmetric equilibrium and beneficial tax havens. Journal of International Economics 81, 253-264.

Johannesen N., 2012. Optimal fiscal barriers to international economic integration in the presence of tax havens. Journal of Public Economics 96, 400-416.

Johannesen N., T. Tørsløv and L. Wier. Are less developed countries more exposed to multinational tax avoidance? WIDER Working Paper 2016/10.

Junquera-Varela R.F., R. Awasthi, O. Balabushko, A. Nurshaikhova, 2019. Thinking strategically about revenue administration reform: The creation of integrated, autonomous revenue bodies, Governance Discussion Paper No. 4. Washington DC: World Bank

Keen M., 2008. VAT, tariffs, and withholding: Border taxes and informality in developing countries. Journal of Public Economics 92, 1892-1906.

Keen M., 2012. Taxation and development – again. IMF Working Paper WP/12/220.

Keen M. and J. Slemrod, 2017. Optimal tax administration. Journal of Public Economics 152, 133-142.

Kleven H.J., C.T. Kreiner and E. Saez, 2016. Why can modern governments tax so much? An agency model of firms as fiscal intermediaries. Economica 83, 219-246.

La Porta R. and A. Shleifer, 2014. Informality and development. The Journal of Economic Perspectives 28, 109-126.

McLaren J. (ed.), 2003. *Institutional elements of tax design and reform.* World Bank Technical Paper No. 539, Washington D.C.

Moore M., O.-H. Fjeldstad, J. Isaksen, O. Lundstøl, R. McCluskey, W. Prichard, 2015. Building tax capacity in developing countries. IDS Policy Briefing 96, Institute of Development Studies, Brighton.

OECD, 1998. Harmful tax competition: an emerging global issue. OECD Publishing.

OECD, 2013. Action plan on base erosion and profit shifting. OECD Publishing.

OECD, 2013. Tax administration 2013: Comparative information on OECD and other advanced and emerging economies. OECD Publishing.

Oguttu A.W., 2016. Tax base erosion and profit shifting in Africa – Part 1: Africa's response to the OECD BEPS action plan. ICTD Working Paper 54.

Sandmo A., 1981. Income tax evasion, labor supply and the equity-efficiency tradeoff. Journal of Public Economics 16, 265-288.

Overesch M., 2009. The effects of multinationals' profit shifting activities on real investments. National Tax Journal 62, 5-23.

Slemrod J., 2001. A general model of the behavioral response to taxation. International Tax and Public Finance 8, 119-128.

Slemrod J. and J.D. Wilson, 2009. Tax competition with parasitic tax havens. Journal of Public Economics 93, 1261-1270.

Slemrod J. and S. Yitzhaki, 2002. *Tax avoidance, evasion, and administration*, in: Auerbach A. and M. Feldstein (eds.), Handbook of Public Economics, 1423-1470, Amsterdam: North-Holland.

Taliercio R.J., 2004. Designing performance: The semi-autonomous revenue authority model in Africa and Latin America, World Bank Policy Research Working Paper No. 3423.

Terkper S.E., 2008. Accounting challenges for semi-autonomous revenue agencies in developing countries. IMF Working Paper WP/08/116.

A Optimal tax system

A.1 Deriving the θ-terms

Totally differentiating the first-order conditions (9) and (10) yields

$$
\begin{aligned}
\theta_1 &= \frac{(1-\alpha)^2[f'(k)-r]^2}{\psi f''(k)} - \frac{\phi}{l} + (1-\alpha)[f'(k)-r]\frac{\partial k}{\partial t} - \phi\frac{\partial \alpha}{\partial t} \\
&- \frac{2t(1-\alpha)[f'(k)-r]^2}{\psi f''(k)}\frac{\partial \alpha}{\partial t} + \frac{2t(1-\alpha)^2[f'(k)-r]f''(k)}{\psi f''(k)}\frac{\partial k}{\partial t} \\
&- \frac{t(1-\alpha)^2[f'(k)-r]^2 f''(k)}{[\psi f''(k)]^2}\frac{\partial \psi}{\partial t} - \frac{t[f'(k)-r]}{l}\frac{\partial k}{\partial t} \\
\theta_2 &= (1-\alpha)[f'(k)-r]\frac{\partial k}{\partial l} - \phi\frac{\partial \alpha}{\partial l} \\
&- \frac{2t(1-\alpha)[f'(k)-r]^2}{\psi f''(k)}\frac{\partial \alpha}{\partial l} + \frac{2t(1-\alpha)^2[f'(k)-r]f''(k)}{\psi f''(k)}\frac{\partial k}{\partial l} \\
&- \frac{t(1-\alpha)^2[f'(k)-r]^2 f''(k)}{[\psi f''(k)]^2}\frac{\partial \psi}{\partial l} - \frac{t[f'(k)-r]}{l}\frac{\partial k}{\partial l} + \frac{t\phi}{l^2} \\
\theta_3 &= \frac{(1-\alpha)\alpha^2[f'(k)-r]^2}{2\psi f''(k)} + \frac{\phi\alpha}{l} - \frac{t\alpha^2[f'(k)-r]^2}{2\psi f''(k)}\frac{\partial \alpha}{\partial t} \\
&+ \frac{t(1-\alpha)\alpha^2[f'(k)-r]f''(k)}{\psi f''(k)}\frac{\partial k}{\partial t} + \frac{t(1-\alpha)\alpha[f'(k)-r]^2}{\psi f''(k)}\frac{\partial \alpha}{\partial t} \\
&- \frac{t(1-\alpha)\alpha^2[f'(k)-r]^2 f''(k)}{2[\psi f''(k)]^2}\frac{\partial \psi}{\partial t} + \frac{t\phi}{l}\frac{\partial \alpha}{\partial t} + \frac{t\alpha[f'(k)-r]}{l}\frac{\partial k}{\partial t} \\
\theta_4 &= -\frac{t\alpha^2[f'(k)-r]^2}{2\psi f''(k)}\frac{\partial \alpha}{\partial l} + \frac{t(1-\alpha)\alpha^2[f'(k)-r]f''(k)}{\psi f''(k)}\frac{\partial k}{\partial l} \\
&+ \frac{t(1-\alpha)\alpha[f'(k)-r]^2}{\psi f''(k)}\frac{\partial \alpha}{\partial l} - \frac{t(1-\alpha)\alpha^2[f'(k)-r]^2 f''(k)}{2[\psi f''(k)]^2}\frac{\partial \psi}{\partial l} \\
&+ \frac{t\alpha[f'(k)-r]}{l}\frac{\partial k}{\partial l} + \frac{t\phi}{l}\frac{\partial \alpha}{\partial l} - \frac{t\alpha\phi}{l^2} - \kappa''(l) \\
\theta_5 &= (1-\alpha)[f'(k)-r]\frac{\partial k}{\partial \mu} + \frac{2t(1-\alpha)^2[f'(k)-r]f''(k)}{\psi f''(k)}\frac{\partial k}{\partial \mu} - \frac{t[f'(k)-r]}{l}\frac{\partial k}{\partial \mu} \\
\theta_6 &= \frac{t(1-\alpha)\alpha^2[f'(k)-r]f''(k)}{\psi f''(k)}\frac{\partial k}{\partial \mu} + \frac{t\alpha[f'(k)-r]}{l}\frac{\partial k}{\partial \mu}
\end{aligned}
$$

Using equations (5) and (7) yields

$$\begin{aligned}
\theta_1 &= \frac{2(1-\alpha)[f'(k)-r]^2}{\psi f''(k)}\left[(1-\alpha)+\frac{3t(1-\alpha)^2}{2\psi}-\frac{3t}{2l}\right]-\frac{2\phi}{l}\\
\theta_2 &= \frac{2\alpha[f'(k)-r]^2}{\psi f''(k)}\left[\frac{(1-\alpha)\alpha}{4}+\frac{t(1-\alpha)}{l}+\frac{3t(1-\alpha)^2\alpha}{4\psi}-\frac{t\alpha}{4l}\right]+\frac{\phi\alpha}{l}+\frac{t\phi}{l^2}\\
\theta_3 &= \frac{2\alpha[f'(k)-r]^2}{\psi f''(k)}\left[\frac{(1-\alpha)\alpha}{4}+\frac{t(1-\alpha)}{l}+\frac{3t(1-\alpha)^2\alpha}{4\psi}-\frac{t\alpha}{4l}\right]+\frac{\phi\alpha}{l}+\frac{t\phi}{l^2}\\
\theta_4 &= \frac{t\alpha^2[f'(k)-r]^2}{\psi f''(k)}\left[\frac{\alpha}{l}-\frac{(1-\alpha)}{l}+\frac{3(1-\alpha)\alpha^2}{4\psi}\right]-\frac{2t\alpha\phi}{l^2}-\kappa''(l)\\
\theta_5 &= -\frac{[f'(k)-r]}{\mu^2\psi f''(k)}\left[(1-\alpha)+\frac{2t(1-\alpha)^2}{\psi}-\frac{t}{l}\right]\\
\theta_6 &= -\frac{t\alpha[f'(k)-r]}{\mu^2\psi f''(k)}\left[\frac{(1-\alpha)\alpha}{\psi}+\frac{1}{l}\right]
\end{aligned}$$

θ_1 and θ_4 are the second-order conditions for t, respectively l, and have to be negative to obtain a maximum. A necessary condition for $\theta_1 < 0$ is that $(1-\alpha)+\frac{3t(1-\alpha)^2}{2\psi}-\frac{3t}{2l} > 0$. This condition implies that $\theta_5 > 0$. Similarly, a necessary condition for $\theta_4 < 0$ is that $\frac{\alpha}{l}-\frac{(1-\alpha)}{l}+\frac{3(1-\alpha)\alpha^2}{4\psi} > 0$.

A.2 Optimal tax policies and development

In what follows, I derive the effects of a change in μ on the tax rate and the tax administration's capacity, which are given by

$$\frac{dt}{d\mu}=\frac{\theta_2\theta_6-\theta_4\theta_5}{|A|}, \qquad \frac{dl}{d\mu}=\frac{\theta_2\theta_5-\theta_1\theta_6}{|A|}, \tag{A.1}$$

where $|A| = \theta_1\theta_4-(\theta_2)^2 > 0$. I distinguish developed and developing countries by the fact that a higher development is associated with a higher investment level and thus a lower marginal net return on investment $f'(k)-r$. In order to draw inferences for developed and developing countries, I evaluate the sign of $\frac{dt}{d\mu}$ and $\frac{dl}{d\mu}$ for a high, respectively a low, level of development.

A.2.1 Developed countries

From the previous analysis, it is clear that $\theta_5 > 0$ and $\theta_6 > 0$. Moreover, because μ is large for developed countries, the return on investment $f'(k) - r$ is low and hence $\theta_2 > 0$. Since $\theta_1 < 0$ and $\theta_4 < 0$, I immediately arrive at $\frac{dt}{d\mu} > 0$ and $\frac{dl}{d\mu} > 0$.

A.2.2 Developing countries

In this part, I will determine the effect of μ on t and l when the initial level of μ is low and hence the return on investment $f'(k) - r$ is high. Because $f'(k) - r$ is high in developing countries, only the terms containing polynomials of $f'(k) - r$ with the highest degree will be relevant for determining the sign of $\frac{dt}{d\mu}$ and $\frac{dl}{d\mu}$.

Tax rate Indicating by β the sum of all terms containing polynomials of $f'(k) - r$ with a lower degree than 3, I get

$$\begin{aligned}
&\theta_2\theta_6 - \theta_4\theta_5 = \beta \\
&- \frac{t\alpha^2[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{(1-\alpha)\alpha}{2} + \frac{2t(1-\alpha)}{l} + \frac{3t(1-\alpha)^2\alpha}{2\psi} - \frac{t\alpha}{2l}\right]\left[\frac{(1-\alpha)\alpha}{\psi} + \frac{1}{l}\right] \\
&+ \frac{t\alpha^2[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{\alpha}{l} - \frac{(1-\alpha)}{l} + \frac{3(1-\alpha)\alpha^2}{4\psi}\right]\left[(1-\alpha) + \frac{2t(1-\alpha)^2}{\psi} - \frac{t}{l}\right]
\end{aligned}$$

Factoring out and collecting terms yields

$$\begin{aligned}
&\theta_2\theta_6 - \theta_4\theta_5 = \beta \\
&- \frac{t\alpha^2[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{2t(1-\alpha)^2}{\psi l} + \frac{t(1-\alpha)}{l^2} + \frac{(1-\alpha)^2}{l} + \frac{t\alpha}{2l^2} + \frac{t(1-\alpha)\alpha^2}{4\psi l}\right] \\
&+ \frac{t\alpha^2[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{(1-\alpha)\alpha}{2l} + \frac{(1-\alpha)^2\alpha^2}{4\psi} + \frac{t(1-\alpha)^2\alpha}{2\psi l}\right].
\end{aligned}$$

Using $\alpha = \frac{t-t_0}{l}$ for the first and second term in the last line delivers

$$\begin{aligned}
&\theta_2\theta_6 - \theta_4\theta_5 = \beta \\
&- \frac{t\alpha^2[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{5t(1-\alpha)^2}{4\psi l} + \frac{t(1-\alpha)}{2l^2} + \frac{(1-\alpha)^2}{l} + \frac{t\alpha}{2l^2} + \frac{t(1-\alpha)\alpha^2}{4\psi l}\right] \\
&- \frac{t\alpha^2[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{t_0(1-\alpha)}{2l^2} + \frac{t_0(1-\alpha)^2\alpha}{4\psi l} + \frac{3t(1-\alpha)^3}{4\psi l}\right].
\end{aligned}$$

Because the term β in the first line can be neglected for determining the sign of $\theta_2\theta_6 - \theta_4\theta_5$, I get that $\theta_2\theta_6 - \theta_4\theta_5 < 0$ and thus $\frac{dt}{d\mu} < 0$.

Capacity of the tax administration Indicating by γ the sum of all terms containing polynomials of $f'(k) - r$ with a lower degree than 3, I get

$$\begin{aligned}
&\theta_2\theta_5 - \theta_1\theta_6 = \gamma \\
&- \frac{2\alpha[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{(1-\alpha)\alpha}{4} + \frac{t(1-\alpha)}{l} + \frac{3t(1-\alpha)^2\alpha}{4\psi} - \frac{t\alpha}{4l}\right]\left[(1-\alpha) + \frac{2t(1-\alpha)^2}{\psi} - \frac{t}{l}\right] \\
&+ \frac{2\alpha[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[t(1-\alpha)^2 + \frac{3t^2(1-\alpha)^3}{2\psi} - \frac{3t^2(1-\alpha)}{2l}\right]\left[\frac{(1-\alpha)\alpha}{\psi} + \frac{1}{l}\right]
\end{aligned}$$

Factoring out and collecting terms yields

$$\begin{aligned}
&\theta_2\theta_5 - \theta_1\theta_6 = \gamma \\
&- \frac{2\alpha[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{(1-\alpha)^2\alpha}{4} + \frac{t^2(1-\alpha)}{2l^2} + \frac{t^2(1-\alpha)^2\alpha}{4\psi l} + \frac{t(1-\alpha)^3\alpha}{4\psi}\right] \\
&- \frac{2\alpha[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{t^2(1-\alpha)^3}{2\psi l} + \frac{t^2\alpha}{4l^2}\right] + \frac{2\alpha[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{t(1-\alpha)\alpha}{2l}\right]
\end{aligned}$$

Using $\alpha = \frac{t-t_0}{l}$ for the last term in the last line delivers

$$\begin{aligned}
&\theta_2\theta_5 - \theta_1\theta_6 = \gamma \\
&- \frac{2\alpha[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{(1-\alpha)^2\alpha}{4} + \frac{t^2(1-\alpha)^2\alpha}{4\psi l} + \frac{t(1-\alpha)^3\alpha}{4\psi} + \frac{t^2(1-\alpha)^3}{2\psi l}\right] \\
&- \frac{2\alpha[f'(k)-r]^3}{\mu^2\psi^2[f''(k)]^2}\left[\frac{t^2\alpha}{4l^2} + \frac{tt_0(1-\alpha)}{2l^2}\right]
\end{aligned}$$

Because the term γ in the first line can be neglected for determining the sign of $\theta_2\theta_5 - \theta_1\theta_6$, I get that $\theta_2\theta_5 - \theta_1\theta_6 < 0$ and thus $\frac{dl}{d\mu} < 0$.

B Tax competition

B.1 Comparative statics for capital investments

The first-order condition for the optimal capital investments of the MNE is given by

$$\psi_i[f'(k_i) - r] = \psi_j[f'(k_j) - r]. \tag{B.1}$$

where $\psi_i = \left[(1-t_i)(1-\alpha_i) + (1-t_0)\alpha_i - l_i\frac{(\alpha_i)^2}{2}\right]$. Total differentiation, using the capital constraint $k_i + k_j = \bar{k}$, yields

$$\begin{aligned}
&\psi_i f''(k_i)dk_i - [f'(k_i) - r]\left[(1-\alpha_i)dt_i + \alpha_i dt_0 + \frac{(\alpha_i)^2}{2}dl_i\right] \\
&= -\psi_j f''(k_j)dk_i - [f'(k_j) - r]\left[(1-\alpha_j)dt_j + \alpha_j dt_0 + \frac{(\alpha_j)^2}{2}dl_j\right] + \psi_j f''(k_j)d\bar{k}.
\end{aligned}$$

Rearranging terms results in

$$\begin{aligned}
dk_i &= \frac{(1-\alpha_i)[f'(k_i) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)}\, dt_i - \frac{(1-\alpha_j)[f'(k_j) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)}\, dt_j \\
&+ \frac{\frac{(\alpha_i)^2}{2}[f'(k_i) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)}\, dl_i - \frac{\frac{(\alpha_j)^2}{2}[f'(k_j) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)}\, dl_j \\
&+ \left[\frac{\alpha_i[f'(k_i) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)} - \frac{\alpha_j[f'(k_j) - r]}{\psi_i f''(k_i) + \psi_j f''(k_j)}\right] dt_0 \\
&+ \frac{\psi_j f''(k_j)}{\psi_i f''(k_i) + \psi_j f''(k_j)}\, d\bar{k}.
\end{aligned} \tag{B.2}$$

From (B.2) follow the investment responses of the MNE as given in (19).

B.2 Total differentiation of the first-order conditions

Totally differentiating the first-order conditions for t_i, l_i t_j and l_j leads to

$$\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_5 & \theta_6 & \theta_7 & \theta_8 \\ \theta_9 & \theta_{10} & \theta_{11} & \theta_{12} \\ \theta_{13} & \theta_{14} & \theta_{15} & \theta_{16} \end{bmatrix} \begin{bmatrix} dt_i \\ dl_i \\ dt_j \\ dl_j \end{bmatrix} = \begin{bmatrix} -\theta_{17} \\ -\theta_{18} \\ -\theta_{19} \\ -\theta_{20} \end{bmatrix} d\bar{k} + \begin{bmatrix} -\theta_{21} \\ -\theta_{22} \\ -\theta_{23} \\ -\theta_{24} \end{bmatrix} dt_0, \qquad \text{where} \tag{B.3}$$

$$\theta_1 = \frac{\partial^2 T_i}{\partial (t_i)^2}, \quad \theta_2 = \frac{\partial^2 T_i}{\partial t_i \partial l_i}, \quad \theta_3 = \frac{\partial^2 T_i}{\partial t_i \partial t_j} \quad \theta_4 = \frac{\partial^2 T_i}{\partial t_i \partial l_j}, \quad \theta_5 = \frac{\partial^2 T_i}{\partial l_i \partial t_i}, \quad \theta_6 = \frac{\partial^2 T_i}{\partial (l_i)^2},$$
$$\theta_7 = \frac{\partial^2 T_i}{\partial l_i \partial t_j}, \quad \theta_8 = \frac{\partial^2 T_i}{\partial l_i \partial l_j}, \quad \theta_9 = \frac{\partial^2 T_j}{\partial t_j \partial t_i}, \quad \theta_{10} = \frac{\partial^2 T_j}{\partial t_j \partial l_i}, \quad \theta_{11} = \frac{\partial^2 T_j}{\partial (t_j)^2}, \quad \theta_{12} = \frac{\partial^2 T_j}{\partial t_j \partial l_j},$$
$$\theta_{13} = \frac{\partial^2 T_j}{\partial l_j \partial t_i}, \quad \theta_{14} = \frac{\partial^2 T_j}{\partial l_j \partial l_i}, \quad \theta_{15} = \frac{\partial^2 T_j}{\partial l_j \partial t_j}, \quad \theta_{16} = \frac{\partial^2 T_j}{\partial (l_j)^2}, \quad \theta_{17} = \frac{\partial^2 T_i}{\partial t_i \partial \bar{k}}, \quad \theta_{18} = \frac{\partial^2 T_i}{\partial l_i \partial \bar{k}},$$
$$\theta_{19} = \frac{\partial^2 T_j}{\partial t_j \partial \bar{k}}, \quad \theta_{20} = \frac{\partial^2 T_j}{\partial l_j \partial \bar{k}}, \quad \theta_{21} = \frac{\partial^2 T_i}{\partial t_i \partial t_0}, \quad \theta_{22} = \frac{\partial^2 T_i}{\partial l_i \partial t_0}, \quad \theta_{23} = \frac{\partial^2 T_j}{\partial t_j \partial t_0}, \quad \theta_{24} = \frac{\partial^2 T_j}{\partial l_j \partial t_0}.$$

Because capital is perfectly mobile, the system stated in equation (B.3) simplifies to

$$\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ \theta_2 & \theta_6 & \theta_7 & \theta_8 \\ \theta_3 & \theta_4 & \theta_1 & \theta_2 \\ \theta_7 & \theta_8 & \theta_2 & \theta_6 \end{bmatrix} \begin{bmatrix} dt_i \\ dl_i \\ dt_j \\ dl_j \end{bmatrix} = \begin{bmatrix} -\theta_{17} \\ -\theta_{18} \\ -\theta_{17} \\ -\theta_{18} \end{bmatrix} d\bar{k} + \begin{bmatrix} -\theta_{21} \\ -\theta_{22} \\ -\theta_{21} \\ -\theta_{22} \end{bmatrix} dt_0, \tag{B.4}$$

where θ-terms are given as follows:

$$\begin{aligned}
\theta_1 = \theta_{11} &= \frac{(1-\alpha)[f'(k)-r]^2}{\psi f''(k)}\left[(1-\alpha)+\frac{3t(1-\alpha)^2}{4\psi}-\frac{3t}{2l}\right]-\frac{2\phi}{l}\\
\theta_2 = \theta_5 &= \frac{\alpha[f'(k)-r]^2}{\psi f''(k)}\left[\frac{(1-\alpha)\alpha}{4}+\frac{t(1-\alpha)}{l}+\frac{3t(1-\alpha)^2\alpha}{8\psi}-\frac{t\alpha}{4l}\right]+\frac{\phi\alpha}{l}+\frac{t\phi}{l^2}\\
\theta_3 &= -\frac{(1-\alpha)[f'(k)-r]^2}{2\psi f''(k)}\left[(1-\alpha)+\frac{t(1-\alpha)^2}{2\psi}-\frac{t}{l}\right]\\
\theta_4 &= -\frac{\alpha^2[f'(k)-r]^2}{4\psi f''(k)}\left[(1-\alpha)+\frac{t(1-\alpha)^2}{2\psi}-\frac{t}{l}\right]\\
\theta_6 = \theta_{16} &= \frac{t\alpha^2[f'(k)-r]^2}{2\psi f''(k)}\left[\frac{\alpha}{l}-\frac{(1-\alpha)}{l}+\frac{3(1-\alpha)\alpha^2}{8\psi}\right]-\frac{2t\alpha\phi}{l^2}-\kappa''(l)\\
\theta_7 &= -\frac{t\alpha(1-\alpha)[f'(k)-r]^2}{2\psi f''(k)}\left[\frac{(1-\alpha)\alpha}{4\psi}+\frac{1}{l}\right]\\
\theta_8 &= -\frac{t\alpha^3[f'(k)-r]^2}{4\psi f''(k)}\left[\frac{(1-\alpha)\alpha}{4\psi}+\frac{1}{l}\right]\\
\theta_{17} = \theta_{19} &= \frac{[f'(k)-r]}{2}\left[(1-\alpha)+\frac{t(1-\alpha)^2}{\psi}-\frac{t}{l}\right]\\
\theta_{18} = \theta_{20} &= \frac{t\alpha[f'(k)-r]}{2}\left[\frac{(1-\alpha)\alpha}{2\psi}+\frac{1}{l}\right]\\
\theta_{21} = \theta_{23} &= \frac{\phi}{l}+\frac{t(1-\alpha)[f'(k)-r]^2}{\psi f''(k)}\left[\frac{1}{l}+\frac{(1-\alpha)\alpha}{2\psi}\right]\\
\theta_{22} = \theta_{24} &= \frac{t\alpha[f'(k)-r]^2}{2\psi f''(k)}\left[\frac{\alpha}{2l}-\frac{(1-\alpha)}{l}+\frac{(1-\alpha)\alpha^2}{2\psi}\right]-\frac{t\phi}{l^2}.
\end{aligned}$$

θ_1 and θ_6 are the second-order conditions for t_i, respectively l_i, and have to be negative to obtain a maximum. A necessary condition for $\theta_1 < 0$ is that $(1-\alpha)+\frac{3t(1-\alpha)^2}{4\psi}-\frac{3t}{2l} > 0$, which implies that $\theta_3 > 0$, $\theta_4 > 0$ and $\theta_{17} > 0$. Similarly, a necessary condition for $\theta_6 < 0$ is that $\frac{\alpha}{l}-\frac{(1-\alpha)}{l}+\frac{3(1-\alpha)\alpha^2}{8\psi} > 0$.

B.3 Optimal tax policies

Applying Cramer's rule to the equation system given in (B.4) yields

$$\frac{dt_i}{d\bar{k}} = \frac{(\theta_4)^2\theta_7\theta_{18}+\theta_1(\theta_8)^2\theta_{17}+(\theta_2)^3\theta_{18}+\theta_3(\theta_6)^2\theta_{17}-\theta_1(\theta_6)^2\theta_{17}-(\theta_2)^2\theta_7\theta_{18}-\theta_3(\theta_8)^2\theta_{17}-\theta_2(\theta_4)^2\theta_{18}}{|A|},$$

$$\frac{dl_i}{d\bar{k}} = \frac{\theta_4(\theta_7)^2\theta_{17} + (\theta_1)^2\theta_8\theta_{18} + (\theta_2)^3\theta_{17} + (\theta_3)^2\theta_6\theta_{18} - (\theta_1)^2\theta_6\theta_{18} - \theta_2(\theta_7)^2\theta_{17} - (\theta_3)^2\theta_8\theta_{18} - (\theta_2)^2\theta_4\theta_{17}}{|A|},$$

where $|A| = (\theta_1)^2(\theta_6)^2 + (\theta_2)^2(\theta_7)^2 + (\theta_3)^2(\theta_8)^2 + (\theta_2)^2(\theta_4)^2 - (\theta_4)^2(\theta_7)^2 - (\theta_1)^2(\theta_8)^2 - (\theta_1)^2(\theta_2)^4 - (\theta_3)^2(\theta_6)^2 > 0$ to obtain a maximum.

B.3.1 Developed countries

Corporate income tax rate Because $[f'(k) - r]$ is low in developed countries, only the terms containing polynomials of $[f'(k) - r]$ with the lowest degree will be relevant for determining the sign of $\frac{dt_i}{d\bar{k}}$. Only $(\theta_2)^3\theta_{18}$ and $\theta_1(\theta_6)^2\theta_{17}$ contain terms with polynomials of $[f'(k) - r]$ of degree 1. Thus,

$$sign\left(\frac{dt_i}{d\bar{k}}\right) = sign\left((\theta_2)^3\theta_{18} - \theta_1(\theta_6)^2\theta_{17}\right),$$

which is positive.

Effectiveness of the tax administration I determine the effect of $\bar{k}$ on l_i in a similar way. Only $(\theta_2)^3\theta_{17}$ and $(\theta_1)^2\theta_6\theta_{18}$ contain terms with polynomials of $[f'(k) - r]$ of degree 1. Thus,

$$sign\left(\frac{dl_i}{d\bar{k}}\right) = sign\left((\theta_2)^3\theta_{17} - (\theta_1)^2\theta_6\theta_{18}\right),$$

which is again positive.

B.3.2 Second-stage analysis

If governments institute a SARA, implying that tax administrations' capacity is determined in the second stage, the system on equations (B.4) for determining the effect of

$\bar{k}$ on l_i simplifies as follows:

$$\begin{bmatrix} \theta_6 & \theta_8 \\ \theta_8 & \theta_6 \end{bmatrix} \begin{bmatrix} dl_i \\ dl_j \end{bmatrix} = \begin{bmatrix} -\theta_{18} \\ -\theta_{18} \end{bmatrix} d\bar{k}. \tag{B.5}$$

Applying Cramer's rule to the equation system given in (B.5) yields

$$\frac{dl_i}{d\bar{k}} = \frac{dl_j}{d\bar{k}} = \frac{dl}{d\bar{k}} = \frac{(\theta_8 - \theta_6)\theta_{18}}{(\theta_6)^2 - (\theta_8)^2} > 0,$$

which is positive because $(\theta_6)^2 - (\theta_8)^2$ has to be positive in order to satisfy the conditions for a maximum. $(\theta_6)^2 - (\theta_8)^2 > 0$ if

$$\theta_6 + \theta_8 < 0, \quad \Leftrightarrow \quad \frac{\alpha}{2l} - \frac{(1-\alpha)}{l} - \frac{(1-\alpha)\alpha}{8\psi} + \frac{3(1-\alpha)\alpha^2}{8\psi} > 0.$$

B.4 Fight against tax havens: second-stage analysis

If governments institute a SARA, the matrix for determining the effect of t_0 on l_i simplifies as follows:

$$\begin{bmatrix} \theta_6 & \theta_8 \\ \theta_8 & \theta_6 \end{bmatrix} \begin{bmatrix} dl_i \\ dl_j \end{bmatrix} = \begin{bmatrix} -\theta_{22} \\ -\theta_{22} \end{bmatrix} dt_0 \tag{B.6}$$

Applying Cramer's rule to the equation system given in (B.6) yields

$$\frac{dl_i}{dt_0} = \frac{dl_j}{dt_0} = \frac{(\theta_8 - \theta_6)\theta_{22}}{(\theta_6)^2 - (\theta_8)^2} < 0.$$

It must be that $(\theta_6)^2 - (\theta_8)^2 > 0$ for a maximum to exist, which is the case if

$$\frac{\alpha}{2l} - \frac{(1-\alpha)}{l} - \frac{(1-\alpha)\alpha}{8\psi} + \frac{3(1-\alpha)\alpha^2}{8\psi} > 0.$$

This condition ensures that $\theta_{22} < 0$ if $\bar{k}$ is sufficiently low, which implies that $\frac{dl_i}{dt_0} < 0$.