The Cost and Benefits of Tax Treaties with Investment Hubs: Findings from Sub-Saharan Africa

Sebastian Beer and Jan Loeprick

September 2018

Abstract

This paper investigates the costs and benefits of concluding double tax treaties with investment hubs. Based on a sample of 41 African economies from 1985–2015, the results suggest that signing treaties with investment hubs is not associated with additional investments; yet, these treaties tend to come with nonnegligible revenue losses. Building on a theoretical model, the paper investigates the role of treaty shopping in driving nominal investment flows and provides indirect evidence for its importance in the sample

JEL Classification Numbers: F21, F23, H25, H26, H32

Keywords: Tax Treaty Policy, Double Tax Treaties, Domestic Resource Mobilization, Mauritius, Sub-Saharan Africa

Author's E-Mail Address: sbeer@imf.org; jloeprick@worldbank.org

I. INTRODUCTION[1]

Over the past decade international taxation has become headline news. Leaked information on private wealth and corporate profits hidden in Switzerland, Luxembourg, Panama, and small island jurisdictions, caused public outcry. Anecdotes of aggressive tax evasion complement macro assessments of the "scourge of tax havens" (Zucman 2015), looking at offshore household wealth (Alstadsaeter and others, 2017) and corporate profit shifting. Yet, despite increasing attention, public scrutiny of policies enabling or facilitating an outsize role of investment hubs[2] in the global economy remains limited,[3] which is likely due to the opacity of the international tax architecture.

In this context, we investigate the role of tax treaties, an aspect of tax policy making that can contribute to the strength of the business model in 'offshore jurisdictions' or 'investment hubs'. Double Tax Treaties (DTTs) are intended to facilitate cross-border trade and investment by demarcating taxing rights of countries. They affect returns to foreign investment by determining when a branch is liable for income tax, the treatment of capital gains, or by changing the taxation of dividend, interest and royalty flows from the country where the investment is made, to the country where it has been initiated. While DTTs serve more purposes and often also come with a range of complementary political objectives, the avoidance of double taxation is at their core. However, these bilaterally negotiated agreements can, by restricting taxing rights and through the abuse of treaty networks using shell companies, also contribute to the loss of a country's tax base and facilitate tax avoidance.

Our paper is focused on tax treaty policy in Sub-Saharan Africa. Building on a panel data set of 41 African economies from 1985 to 2015, we examine the effects of concluding DTTs with investment hubs empirically. In Africa, Mauritius has positioned itself as a conduit country for investment into the continent. As part of its strategy, Mauritius taxes corporate income at 15 percent, which can be reduced to an effective rate of 3 percent,[4] and negotiated favorable tax treaties with many African economies. We take advantage of the fact that 11

[1] The authors would like to thank Geerten Michielse for his guidance and contributions throughout the preparation of this draft. The authors would also like to thank Punam Chuhan-Pole, Matt Collin, Alexander Klemm, Ruud De Mooij, Jeffrey Owens, Victoria Perry, Marijn Verhoeven, and participants of the Vienna University of Economics (WU) research seminar on the future of international taxation for providing helpful comments. The views expressed in this paper are those of the authors and do not necessarily represent the views of the IMF, the World Bank, their Executive Boards, or management.

[2] Definitions of "investment hubs" or "tax havens" differ across institutions and over time and there is no single globally accepted definition. The OECD has developed a list of havens as part of its work on harmful tax practices in 1998 and has since revised the list and approaches to define havens. Other efforts at defining the concept have been made by researchers, legislators, CSOs and regional and international organizations. A close link also exists to the discussion and definition of offshore financial centers (OFC), providing financial services to non-residents.

[3] Several CSOs, including the Tax Justice Network (2015), Oxfam, ActionAid (2016) are publishing lists of tax havens, tax treaty policy decisions and examples of cost. In the G20/OECD led Base Erosion and Profit Shifting (BEPS) process (2015) several measures that will help reduce the risks of tax treaty abuse have been agreed.

[4] See section II. for details.

countries in the region have concluded treaties with Mauritius, but others have not.[5] Using a Difference-in-Difference framework, we do not find that these treaties increase Foreign Direct Investment (FDI) in treaty partner countries. Our work, however, does provide an indication of non-negligible revenue losses in source countries following the enactment of a treaty with Mauritius and investment hubs more generally. These results are consistent with the idea that treaty-conclusion triggers a rerouting of investment and income flows, and potentially increases incentives for profit shifting,[6] rather than increasing the overall investments made.

Subsequently, we test the importance of treaty-shopping in driving these results. Due to lack of bilateral FDI data for the region, which would enable a more direct approach, we employ a novel method to identify the abuse of a country's treaty network using aggregate data. The identification strategy draws on two observations: First, tax avoidance has a negative effect on tax collections and, by reducing MNEs' cost of capital, a positive effect on the optimal investment magnitude. Second, a preferential treatment of some investors, commonly resulting from DTTs, provides tax avoidance opportunities. Accordingly, we identify treaty-shopping by measuring the systematic response of investment and government revenues to variation in a source country's tax treatment of cross-border income flows. The results of our non-linear system estimations indicate that treaty-shopping drives nominal investment flows and reduces tax revenue across the region. Moreover, we find that treaty-shopping is more pervasive in countries that have signed DTTs with investment hubs. Back-of-the-envelope calculations indicate that revenue losses due to treaty-shopping amount to 5 percent of CIT revenue across SSA. In countries that have concluded a DTT with investment hubs these costs increase on average to 15 percent of CIT revenue.

These estimates for African economies are surprisingly large and deviate from common assessments of global revenue risks linked to aggressive avoidance and evasion (see e.g. Beer and others, 2018). Several factors potentially explain this finding. First, prior studies typically identify international tax avoidance by measuring the systematic response to CIT rate differentials using firm-level data. This strategy does not capture treaty-shopping incentives, which can play a significant role in reducing a group's effective tax burden. Second, we quantify effects for SSA, a region which is potentially more vulnerable to tax avoidance by MNEs than developed economies (Crivelli and others, 2015). Third, tax avoidance estimates based on macro-data tend to be larger (Beer and others, 2018), either because macro-data captures long-run responses or because of statistical issues with micro- or macro-data.[7] Finally, we measure marginal effects of treaty-conclusion while a country's

[5] Between 1985 and 2015, Mauritius concluded 15 DTTs with economies in SSA. Due to data limitations, we evaluate the effect of 11 treaties.

[6] Unless double tax treaties stipulate otherwise, withholding taxes are typically levied on the gross income payments to foreign entities thus limiting the incentive to engage in tax arbitrage. Where administrative capacity to address profit shifting is limited, withholding taxation can therefore act as 'a second line of defense' in safeguarding the domestic tax base (Balabushko and others, 2017).

[7] Estimates based on macro-data are potentially biased, due to the inability to control for confounding factors, such as firm-level assets, when measuring tax minimization responses. However, also micro-data might be biased if coverage of firm-level data is limited. For instance, where no or incomplete accounting information is available, this could be correlated to more aggressive tax optimization. Given the limited coverage of firm-level

treaty-network likely exerts important non-linear effects on a country's tax revenue. Given the uncertainty about which of these factors drives measured effects, our results need to be interpreted with caution.

Our work adds to a growing, albeit inconclusive, empirical literature on the effect of host country tax regimes and tax treaties on FDI. Egger and others (2006) use propensity score matching and find a negative effect of treaties on outward FDI stocks in the OECD. Lejour (2014) and Barthel and others (2010) find that treaties do increase bilateral FDI, whereas Neumayer (2007) only finds positive effects linked to the conclusion of tax treaties with the United States in middle-income countries. Focusing on Sub-Saharan Africa, our findings suggest that treaties are not associated with increases in foreign direct investment in low-income economies. Inconclusive empirical findings on the importance of tax treaties for FDI are likely partly a result of the subordinated role of taxation in investor decision making,[8] and of the complexity of FDI decision making in multinational enterprises, which are pursuing different strategies (Carr and others, 2001, Bergstrand and Egger, 2007), with different drivers among those seeking to take over foreign assets as opposed to making greenfield investments (Head and Ries, 2008).

Furthermore, our finding that treaty-conclusion with investment hubs primarily leads to rerouting of investment and income flows, is aligned with earlier work on treaty shopping in the Netherlands and Germany. Using micro-data on Dutch Special Purpose Entities, Weyzig's (2013) analysis suggests that reduced withholding tax rates on dividend flows are a driver of FDI diversion via the Netherlands. Similarly, Weichenrieder and Mintz (2008) show that an increase in bilateral withholding taxes to and from Germany increases the probability that outward and inward FDI is diverted via a third country. Petkova and others (2018) control for treaty shopping opportunities in their assessment of the relevance of tax treaties for investment when adopting a network approach that builds on van't Ried and Lejour (2018) and Hong (2018). Noticeably, they find that only the treaties which reduce tax obligations result in higher FDI, thus underscoring the potential importance of treaty shopping in driving treaty effects.

By quantifying the revenue implications of treaty-shopping, we add to recent assessments of the costs associated with tax treaties. Country-specific estimates have, for instance, been provided for Dutch treaties (McGauran, 2013), and the Ukraine (Balabushko et al, 2017). More recently, Jansky and Svidivy (2018) find that costs in terms of revenue foregone amount to up to 0.2 percent of GDP in some countries. Our assessment of the non-negligible cost of tax treaties concluded with investment hubs also raises more general questions on the decision-making processes in countries that are giving up some of their taxing rights. Looking at results of tax treaty negotiations, Hearson (2017) builds on earlier work by Rixen and Schwartz (2009) and Barthel and Neumeyer (2012) and finds that outcomes differ

information and bilateral transaction information for Sub-Saharan African countries, our analysis is based on macroeconomic aggregate information, which is available for most countries, but tends to be less accurate when aiming at identifying tax driven effects.

[8] A recently concluded global investor survey ranks taxation secondary to other factors such as political stability, legal and regulatory environment, and market size (WBG 2017).

noticeably across simpler and more complex provisions, and, relatedly, demonstrates a learning effect: Developing countries seem to maintain more taxing rights in more recent treaties in their network.[9]

Moreover, our work adds to the recent strand of literature that focuses on the interaction between tax avoidance and real activity. De Mooij and Liu (2018) assess the effect of transfer-pricing regulation on real investment. Using firm-level data, they find that tightening anti-avoidance measures can reduce investment activity. Buettner and others (2017) also find negative effects of introducing thin capitalization rules on FDI and employment and observe a higher tax rate sensitivity of FDI in countries with tighter measures targeting debt-shifting. Similarly, our estimations indicate that treaty-shopping opportunities exert a small, positive effect on FDI. Stricter substance requirements in DTTs may increase MNE's cost of capital, thus reducing the optimal investment magnitude at the margin. Finally, we improve on previous identification strategies of tax avoidance[10] by exploiting the joint distribution of two variables, FDI and revenue, drawing on our theoretical model.

The paper proceeds as follows. Section II provides background on DTTs in SSA and Mauritius' tax policy regime. We then present the theoretical underpinnings of our identification strategy (Section III), followed by a discussion of our empirical approaches and descriptive statistics (Section IV). Results on the impact of concluding a treaty with Mauritius and for our estimation of the structural model are presented in Section V. Section VI concludes.

II. DTTs AND THE ROLE OF 'INVESTMENT HUBS' IN SSA

DTTs are aimed at avoiding a situation where the same income is taxed twice by two different jurisdictions. By defining when and what type of a taxpayer's activity a country can tax, DTTs provide legal certainty for cross-border investments, and are therefore expected to foster FDI and, more generally, the economic ties between treaty partners. Today, more than 3000 DTTs exist, most of them between developed economies.

A central component of treaties is the definition of the withholding tax rates a source country can apply on cross-border income payments, such as royalties, interest or service fees. Between high-tax countries, limitations on withholding taxation tend to result in a transfer of tax revenue between governments, as foreign withholding tax payments are typically creditable against the domestic tax liability in resident countries. In contrast, the limitation of withholding tax rates on payments directed toward low-tax jurisdictions results in a transfer from governments to taxpayers, as any tax payment would not have been (fully) creditable in the low-tax jurisdiction. Notably, DTTs are bilateral agreements aimed at facilitating investments and trade between residents of two countries. Residents of third countries are explicitly excluded from treaty benefits. However, definitions of treaty beneficiaries differ,

[9] Interestingly, Braun and Zagler's (2018) analysis suggests that, in the context of developing countries, foreign aid may sometimes be an instrument to compensate for asymmetries in their tax treaties.

[10] See Beer and others (2018) for a review of the empirical literature on international tax avoidance by MNEs.

and weak substance requirement can imply that "any one treaty should be considered a potential treaty with the world."[11],[12]

Table 1. Important Treaty-Partner Countries for SSA

Country	GDP (in billion US Dollar)	Number of effective treaties with SSA	Average WHT rate with SSA (in %)	Domestic CIT rate (in %)
Mauritius	12.17	15	6.83	15 (3)
South Africa	295.46	15	9.97	28
Canada	1529.76	11	14.93	26.5
France	2465.45	11	10.68	33.31
Belgium	467.96	10	10.30	33.99

Notes: table summarizes key features of treaty partner countries for SSA, ordered by the number of concluded treaties between 1985 and 2016. Data sources: IBFD, IMF

Table 1 summarizes the main treaty partner countries for SSA economies, ranked by the number of treaties concluded between 1985 and 2016. Most of the listed countries provide a large pool of potential resident investors and apply non-negligible tax rates on domestic income. Mauritius, however, has a comparatively small base of potential domestic investors and taxes income minimally under its global business companies (GBC) regime: if treated as a resident, GBCs are subject to the standard corporate income tax rate of 15 percent, but can utilize a unilateral foreign tax credit which effectively reduces the Mauritian corporate income tax burden on foreign source income to 3 percent.[13]

Mauritius' importance for SSA stems from its role as an investment hub. Corporate groups often include intermediate holding companies, which can be held for various purposes, such as the coordination and centralization of certain group activities (typically factoring and insurance), group financing, or holding intellectual property. Another objective is often the reduction of an MNE's tax liability. To attract these intermediate holding companies, investment hubs provide three related features: First, they have an extensive tax treaty network, ideally containing treaties with low or zero withholding tax rates (and other tax benefits) with source countries. At the same time, they do not levy any withholding tax on the outflow of income. Second, investment hubs have 'liberal' substance requirements and although it is not necessary to maintain strict non-disclosure rules (bank secrecy), recent leaks suggest that such rules tend to help generate business. Third, investment hubs tax income minimally. Low or

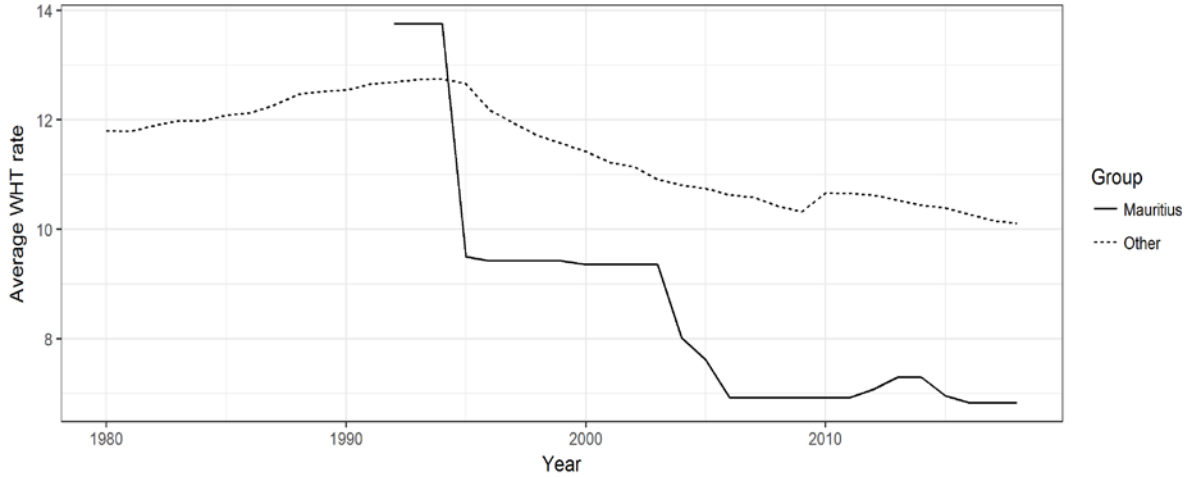
[11] See 2016, blog post by Jim Brumby and Michael Keen referencing a presentation by Stephen Shay: http://blogs.worldbank.org/governance/tax-treaties-boost-or-bane-development.

[12] In recent years, the role of tax treaties has been reevaluated, most notably in the context of work led by the OECD aimed at avoiding or countering the abuse of tax treaty provisions. Treaty anti-abuse provisions to counter treaty shopping comprise a minimum standard under the OECD/G20 BEPS initiative. More specifically, the minimum standard requires: (i) inclusion of a clear statement in tax treaties of the intention to avoid non-taxation or reduced taxation through tax evasion or avoidance, including through treaty shopping; and (ii) the adoption of one or both of a limitation of benefits (LOB) rule or a principal purpose test (PPT) rule, supplemented as needed to deal with conduit arrangements.

[13] The allowable business activities of a GBC (category 1) include, among others, assets management, consultancy services, financial services, insurance, logistics, or marketing. All those activities are non-location specific and thus easily transferable within any MNE. A GBC is regarded as tax resident in Mauritius if they can demonstrate that their management and control is situated in Mauritius.

zero income tax rates in investment hubs are usually achieved by offering specific preferential regimes. These can be created as an offshore regime that is applicable to resident companies owned by nonresidents, but also as a special tax regime for certain activities.

Figure 1. Average WHT Rates for Dividends, Interests, and Royalties, in Double Tax Treaties signed by Sub-Saharan African Economies with the World and Mauritius



Data Source: IBFD 2018

Starting in the 1990s, Mauritius expanded its treaty network with both low-tax jurisdictions and African economies.[14] Of Mauritius' 43 effective treaties today, around 20 percent have been concluded with low-tax jurisdictions and 35 percent have been signed with African economies. Figure 1 contrasts the average withholding tax rate between African economies in our sample with Mauritius and with other economies, showing that Mauritius managed to negotiate favorable terms in many of its treaties with SSA. While the first treaty concluded by Mauritius in the region (Zimbabwe, 1992) included relatively high withholding tax rates, subsequent treaties came with much lower rates. In addition, Mauritius does not levy withholding taxes on outgoing interest or dividend payments. The extensive treaty network with both low-tax jurisdictions and African economies in combination with no domestic withholding taxation and an effective CIT rate of 3 percent provides for an easy flow of capital via Mauritius, which neatly fits the definition of an investment hub. Several reforms to the taxation of the Mauritian global business sector were announced recently, which are aimed at implementing international standards targeting harmful tax practices.[15]

[14] Mauritius signed its first treaties with EU countries in the early 1980s. Starting in the 1990s, Mauritius expanded its network with other investment hubs, including Barbados, Cyprus, Malta, Luxembourg, Guernsey, Monaco, the Seychelles, and Singapore, and with African countries, thus enhancing its attractiveness as an investment hub for the region. In addition to the 43 effective treaties, 31 DTTs are being negotiated, signed, or are awaiting ratification. An important initial driver of the growth of Mauritius's offshore sector was a double tax treaty signed with India in 1982 that was recently renegotiated to gradually remove its capital gains tax exemption from 2017-2019.

[15] With the most recent budget (June 14, 2018), a range of policy changes for the Global Business license regime were announced. These include a removal of the deemed foreign tax credit for Category 1 license holders and stronger substance requirements, which will be determined by the Financial Services Commission.

Under the G20/OECD Base Erosion and Profit Shifting (BEPS) initiative, there is also ongoing pressure to improve some of the anti-abuse provisions in Mauritius' tax treaties. Mauritius is a member of the BEPS Inclusive Framework and has therefore committed to implement the BEPS minimum standards. Further, Mauritius signed the Multilateral Convention (MLI) on July 5, 2017 which will accelerate adoption of anti-abuse rules. However, core provisions of concern for many of its treaty partners are the withholding rates agreed on dividend, interest and royalty payments, which are not (yet) covered under these initiatives.[16]

To evaluate the effects of DTTs with investment hubs, we initially focus our paper on treaties concluded with Mauritius for several related reasons. First, Mauritius is the most important offshore center for the region. While other investment hubs have also concluded DTTs with SSA, their regional treaty networks are less comprehensive.[17] Second, many of Mauritius' tax treaties are recent. Given that the patchwork of African tax treaties has been constructed over more than 50 years in changing economic and political contexts (Hearson 2015), the focus on Mauritius allows us to focus on contemporary policy decisions. Third, Mauritius' treaty network expanded extensively over a relatively short time frame. This provides the necessary variation to identify the impact of those treaties empirically. Fourth, the role of Mauritius and the cost-benefit of its treaty links to SSA has been receiving increasing attention in the past couple of years. South Africa[18] and Rwanda[19] recently renegotiated their treaties with Mauritius to address imbalances. In Kenya, the Tax Justice Network (TJN), a prominent NGO focusing on international tax issues, has drawn attention to the signature of a tax treaty with Mauritius (2012) by bringing a case against the agreement to the Kenyan High Court in 2014.[20] And, in November 2017, the publication of the Paradise Papers led to additional scrutiny of the cost of Mauritius positioning itself as the investment hub of choice for investors in Africa and raised further questions on the additionality of investments made via

[16] Mauritius opted for bilateral discussions with various treaty partners, thus allowing for discussions on all aspects of treaty policy, including provisions not covered by the MLI.

[17] As an extension, we also assess the effect of treaty conclusion with investment hubs more generally, by redefining our treatment variable to include DTTs between SSA economies and countries where the sum of inward and outward FDI stocks exceeds two times the domestic GDP in 2016 (for a similar definition see Daamgard and Elkjaer, 2017). The extended list of countries meeting this definition includes Malta; Cyprus; Hong Kong SAR, China; Luxembourg; Singapore; Ireland; Switzerland; the Netherlands; Barbados; and the Seychelles. However, as none of these countries has concluded more than 5 treaties within our sample period in the region, they do not lend themselves for a separate empirical evaluation.

[18] The amendments included among others the introduction of withholding tax on interest earned in South Africa and introduced a new tie-breaker clause to avoid double non-taxation. Withholding rates in the treaty remain, however, more attractive than the regular domestic rate. See Davis Tax Committee report on preventing tax treaty abuse, p.73-76; http://www.taxcom.org.za/docs/New_Folder3/8%20BEPS%20Final%20Report%20-%20Action%206.pdf.

[19] In 2013, with significant increases in withholding rates among others.

[20] TJN argued that the treaty violates the Kenyan constitution by going against principles of good governance, sustainability and accountability.

Mauritius.[21] Yet, to date, the effects of the Mauritian treaty network have not been systematically quantified.

III. CONCEPTUAL FRAMEWORK

We formalize the basic cost-benefit trade-off decision makers face when concluding DTTs next. In our model, we consider an MNE which may finance an investment project in a source country from N different overseas locations (resident countries). We consider three decision margins: (i.) the total investment magnitude, (ii.) the funding share provided by each location, and (iii.) the overall financing mix of the investment. We model the effect of DTTs through their impact on bilateral withholding tax rates. Equations summarizing the firm's optimal investment decision and the implied tax revenue are the basis of our empirical strategy in the next section.

A. The Optimal Investment Magnitude and Rerouting of Income

Let the concave production function $f(K)$ represent the project's pre-tax return from investing K units of capital in source country S. The investment is jointly financed from N overseas locations, where location $i = 1, \dots N$ provides a share α_i of the total capital. We ignore transfer pricing issues by assuming that the firm in the source country remunerates its overseas affiliates at marginal financing costs, *r*, and in proportion to the respective capital injection. The after-tax return of the investment project can then be written as

$$\pi = f(K)(1 - t_s) - \left(1 + \sum_i \alpha_i \tau_i\right) rK \qquad (1)$$

where t_s is the corporate tax rate in the source country and τ_i is the tax costs of income repatriation from source country S to location i.[22] We derive a general expression for the tax cost of income repatriation in the following section and show that τ_i can turn negative where DTTs reduce withholding tax charges substantially. Without further restrictions, the MNE is incentivized to finance the entire project from a single location with the lowest tax cost of income repatriation.

[21] The International Consortium of Investigative Journalists (ICIJ), for instance, points toward structures that have been used to shift profits out of African economies toward Mauritian holding companies with limited or no meaningful functions (see ICIJ, 2017). The Mauritian Economic Development Board's official response submits that the tax advantages accorded were required for the viability of the project in question (see: EDB, 2018).

[22] For instance, if affiliates in location *i* provide debt and no withholding taxes apply on interest payments to this location, the effective tax cost of income repatriation would read $\tau_i = t_i - t_s$, where t_i denotes the corporate income tax rate in country i.So, if the entire project was financed through this location, the after-tax return would simplify to $\pi = [f(K) - rK](1 - t_s) + rK(1 - t_i) - rK$, where the first term represents corporate profit in the source country, the second term is profit in the resident country, and the third term represents (non-deductible) financing costs in the resident country.

However, source countries' FDI is not originating from a single resident location in practice. To reflect this fact in our model, we incorporate costs to treaty-shopping. Specifically, we assume that the MNE incurs costs C_i if the share of funds provided from location i deviates from this location's actual funding potential $\overline{\alpha_i}$. In the empirical application, we use GDP weights to proxy for funding potential. Following the literature on international profit shifting (see e.g. Huizinga and Laeven,2008), we assume treaty shopping costs take the simple quadratic functional form:

$$C_i = \frac{N}{2\gamma}(\alpha_i - \overline{\alpha_i})^2 rK, \qquad (2)$$

where the parameter γ determines the overall costs of treaty shopping.[23] A deviation between real funding potential and reported financing is costly as it requires the establishment of new organizational structures, which may need to have sufficient substance for an entity to qualify for treaty benefits.[24] As $\gamma \to 0$, treaty shopping costs tend to infinity, in which case the actual funding share in each location is no longer a choice variable.

Equation (2) suggests that *marginal* tax avoidance costs are linear in the avoided tax base. In the context of our model, this implies that variation in a country's bilateral tax treatment of income flows provides an identification device for treaty-shopping. In contrast, van't Ried and Lejour (2018) employ a network analysis and assume that *total* tax avoidance costs are linear in the avoided tax base. As a result, the minimal bilateral tax burden determines income flows in their framework, rather than the entire distribution of a country's bilateral tax rates. Note that all investments should be rerouted through a single conduit country if total treaty-shopping costs are linear in the avoided tax base. This assumption is easier to reconcile with empirical data when using bilateral FDI data, such as van't Ried and Lejour (2018), than when using aggregate FDI data as we do.[25]

The MNE seeks to maximize its global return after-tax and after treaty-shopping by choosing the investment magnitude and the place of financing subject to the constraint that the investment is fully financed $\sum_i \alpha_i = 1$. Formally, the MNE's maximization problem reads:

$$\max_{K,\beta} L = \pi - \sum_i C_i + \lambda \left[\sum_i \alpha_i - 1 \right]$$

[23] For simplicity, we assume in the theoretical model that these costs are constant across all resident locations. In the empirical section, however, we allow them to vary between investment hubs and other locations.

[24] Arguably costs of treaty shopping have escalated due to the recent emphasis in the OECD/G20 BEPS initiative on introducing anti-abuse provisions that require sufficient substance for an entity to qualify for treaty benefits.

[25] In other words, it is easier to argue that all investments from country P to country S are rerouted through conduit country C, than to argue that all investments in country S are structured through country C.

where λ is the Lagrangian multiplier associated with the equality-constraint. Necessary and sufficient[26] conditions are

$$\frac{\partial L}{\partial \beta_i} = -rK\tau_i - \frac{N}{\gamma}(\alpha_i - \overline{\alpha_i}) + \lambda = 0 \qquad for\ \ i = 1, \dots N \tag{3.a}$$

$$\frac{\partial L}{\partial K} = f'(K)(1 - t_s) - \left(1 + \sum_i \alpha_i \tau_i\right) r - \frac{N}{2\gamma}\left(\beta_i - \overline{\beta_i}\right)^2 = 0 \tag{3.b}$$

$$\frac{\partial L}{\partial \lambda} = \sum_i \alpha_i - 1 = 0 \tag{3.c}$$

Combining (3.a) and (3.c) describes the optimal financing scheme as

$$\alpha_i = \overline{\alpha_i} - \gamma \frac{\tau_i - \bar{\tau}}{N}, \qquad for\ \ i = 1, \dots N \tag{4}$$

where $\bar{\tau} = \frac{1}{N}\sum_i \tau_i$ is the average tax cost of income repatriation. Affiliates in location *i* thus provide less than their actual funding potential ($\alpha_i < \overline{\alpha_i}$) if the tax cost of repatriating income to this location is above average. Conversely, they provide a larger share of the total investment than their funding potential ($\alpha_i > \overline{\alpha_i}$) if the tax cost of income repatriation to this location are below average. The observed funding share in each location coincides with the hypothetical funding potential once treaty-shopping becomes excessively expensive ($\gamma \to 0$). Similarly, these two measures correspond if the MNE's tax cost of income repatriation are constant across locations.

Next, substituting the optimal financing decision (4) in (3.b), we find

$$f'(K) = \frac{r}{1 - t_s}\left[1 + E(\tau_i) - \gamma \frac{1}{2} Var(\tau_i)\right], \tag{5}$$

where $E(\tau_i) = \sum_i \overline{\alpha_i}\tau_i$, is the funding potential weighted-average tax cost of income repatriation and $Var(\tau_i) = \frac{1}{N}\sum_i (\tau_i - \bar{\tau})^2$ is the variance of the tax cost of income repatriation.

Equation (5) implicitly defines the optimal investment magnitude, by requiring the marginal pre-tax return on the investment matches the user cost of capital. In contrast to more traditional representations, the user cost of capital in our model incorporates the effect of cross-border taxation and treaty-shopping. It clarifies that, by diminishing after-tax returns, tax costs of income repatriation weigh negatively on the optimal investment magnitude.

[26] Sufficiency follows from concavity of the production function. Proofs are available upon request.

Treaty-shopping reduces the MNE's financing costs, with positive effects on the optimal investment magnitude.[27]

B. DTTs and the Tax Costs of Income Repatriation

We next derive the tax costs of income repatriation for equity and debt financed investments to clarify the role of DTTs and the relevance of the withholding taxation system for investment decisions.

Dividend payments are generally not deductible from the corporate tax base in source countries and not taxable in most overseas locations. They do, however, trigger withholding taxes at rates which differ depending on the recipient location (i.e. the resident country receiving the dividend). Denote with ω_i^{div} the withholding tax rate charged by the source country on dividend payments directed to country i. The effective withholding tax rate is the lower of two rates: the withholding tax rate specified in the domestic law and the maximal withholding tax rate agreed upon in the DTT with location i. The tax cost of repatriating one unit of after-tax income in the form of dividends is then simply[28] $\tau_i^{equity} \equiv \omega_i^{div}$.

For interest income, cross-border interest payments are generally deductible from the corporate tax base in source countries and taxable in investor locations. In addition, interest payments also trigger recipient-specific withholding taxes in the source country. To avoid double taxation, we assume that all locations provide a tax credit for foreign withholding tax payments which is capped at the domestic tax liability. Let the corporate tax rate in the source and investor country *i* be given by t_s and t_i, respectively and denote the withholding tax rate on interest payments paid to location *i* by ω_i^{int}. The tax cost of repatriating one unit of after-tax income in the form of interest payments to location i is then given by $\tau_i^{debt} \equiv t_i - t_s$ if the tax liability in the investor country is large enough to make use of the tax credit in full, i.e., $t_i > \omega_i^{int}$ and $\omega_i^{int} - t_s$ otherwise.

We assume the investment project is mutually financed from all locations, each of which provides debt whenever the tax consequences of doing so are preferential.[29] The above definitions imply that debt-financed investments trigger a smaller tax liability, $\tau_i^{debt} < \tau_i^{equity}$, whenever $t_i < t_s + \omega_i^{div} \equiv t_i^H$. The effective tax costs of income repatriation from the source country to affiliate i thus reads

---

[27] For another context where avoidance opportunities have a positive real effect, see e.g. Grubert and Slemrod (1998).

[28] If location i operated a worldwide tax system, we would need to distinguish two cases. First, if the tax credit in the resident country was too small, the tax cost of income repatriations would remain at the same value ω_i^{div}. With sufficient tax credit in the resident country, the tax cost of income repatriation in the form of dividends would change to $\frac{t_i - t_s}{1 - t_s}$; a grossed-up tax differential.

[29] So, while we model the financing decision in each resident location as a discrete choice, the project is ultimately financed with both financing sources. We therefore also neglect thin-capitalization rules.

$$\tau_i \equiv \begin{cases} \omega_i^{int} - t_s & if \quad t_i \leq \omega_i^{int} \\ t_i - t_s & if \quad \omega_i^{int} < t_i \leq t_i^H \\ \omega_i^{div} & if \quad t_i^H < t_i \end{cases} \tag{6.a}$$

C. Tax Revenue from FDI

The MNE's financing choices also affect government revenues. Debt financed investments from location i raise withholding taxes at a rate of ω_i^{pas} but reduce the corporate tax base at a rate of t_s. In contrast, equity financed investments from this location increase tax revenue by ω_i^{div}. Accordingly, we define the effective tax rate on cross-border income flows to location i by

$$\eta_i \equiv \begin{cases} \omega_i^{int} - t_s \ \ if \ \ t_i \leq t_i^H \\ \omega_i^{div} \ else \end{cases} \tag{6.b}$$

Total tax revenue from foreign investments are the sum of corporate tax revenue $t_s f(K)$ and additional taxation of cross-border income flows. Using definition (6.b) and the optimal funding share as defined in equation (4), we can write total tax revenue from foreign investments as

$$T = t_s f(K) + rK[E(\eta_i) - \gamma Cov(\tau_i, \eta_i)]. \tag{7}$$

Tax revenue from FDI is thus an increasing function of the average effective tax rate on international income flows $E(\eta_i) = \sum_i \bar{\alpha}_i \eta_i$ and the corporate income tax rate, holding the stock of foreign investments constant. Due to treaty-shopping opportunities, however, tax revenue is a decreasing function of the effective tax rate's dispersion $Cov(\tau_i, \eta_i) = \frac{1}{N}\sum_i \eta_i(\tau_i - \bar{\tau})$.

D. Implications for an Optimal Withholding Tax Policy

Equations (5) and (7) formalize the basic cost-benefit trade-off policy makers face when deciding on an appropriate treatment of international income flows: while lower withholding tax charges directly reduce tax collections on existing income flows, they may attract additional investment and thus impact positively on government revenue. Treaty-shopping opportunities alter this trade-off as they provide the MNE with an additional decision margin to respond to policy changes. In effect, this additional instrument allows the MNE to separate tax minimization from production decisions.

We establish this fact formally in Annex 1, when deriving a revenue-maximizing withholding tax rate on passive income flows toward low-tax jurisdictions. Specifically, we show that a maximum of equation (7), where DTTs effectively reduce withholding taxes toward this location ($\tau_i^* < \bar{\tau}$), is characterized by

$$\tau_i^* = \bar{\tau}_{-i} - h(\gamma). \tag{8}$$

Here, $\bar{\tau}_{-i}$ is the average tax cost of income repatriation when excluding location i and h is a positive valued function, which is decreasing in γ.

This solution formalizes an intuitive result: where it is revenue maximizing to reduce withholding tax rates on transactions with a country, there is less justification to do so in the presence of treaty-shopping. In effect, a revenue-maximizing withholding tax policy disallows the deductibility of interest payments as treaty-shopping becomes excessive.[30] This result resembles earlier findings by Best and others (2015), showing that tax avoidance justifies, to some extent, the taxation on a gross rather than on a net basis to support revenue efficiency. While the model is overly simplistic in disregarding other policy objectives, such as employment or wages, it provides a useful framework to understand the importance of treaty-shopping in evaluating a country's treaty policy.

Notably, the model highlights the role of the withholding tax system for MNEs' treaty-shopping opportunities while neglecting potential effects on other avoidance margins. However, withholding taxes likely also affect MNEs' profit-shifting incentives. For instance, tax savings in the source country associated with an internal loan, financed from an affiliated company in a low-tax jurisdiction, increase as the relevant DTT reduces withholding taxes on cross-border interest payments. Lower withholding taxes thus incentive the MNE to charge higher interest payments from the source country affiliate. It follows that the revenue implications from reducing withholding taxes vis-à-vis low-tax jurisdictions likely exceed the effects considered in our model.

IV. Empirical Strategy and Descriptive Statistics

We proceed in two steps to quantify the effects of DTTs. We first assess the overall cost-benefit trade-off inherent in treaty conclusion with investment hubs by taking advantage of the fact that 11 economies in SSA have signed a tax treaty with Mauritius in our sample period, while numerous African economies with similar characteristics have not. This allows us to identify the effect of treaty conclusion with Mauritius in a Difference-in-Difference (DD) framework, comparing a treatment and control group over time. The approach captures the total revenue effects of DTTs concluded with Mauritius, including indirect effects linked to an adaption of internal transfer prices.

While being intuitive, the DD approach comes with several drawbacks. Notably, treaties with Mauritius vary widely with respect to the agreed withholding tax rates and other relevant provisions determining taxing rights. A binary indicator may thus confound different effects. Moreover, the effect of a treaty tends to be related to other treaties in a country's network, complicating the definition on an appropriate counterfactual to treaty-conclusion with a specific country. Building on our theoretical framework, we employ a second estimation strategy to test the effects of DTTs and to quantify the extent of treaty shopping. Specifically, the optimal investment decision (equation 4) and the resulting tax revenue (equation 7)

[30] At the limit, where treaty-shopping cost tend to zero, the tax cost of income repatriation to location i needs to equal the average tax cost of income repatriation to any other location, implying that $\omega_i^{*int} = t_s + \bar{\tau}_{-i}$.

provide two independent identification devices to estimate the hypothetical treaty-shopping cost parameter. By combining this estimate with the data, we delineate the implications of treaty-shopping for investment and government revenue in source countries.

A. Estimation of Total Effects

Direct evidence on the effect of DTT builds on the following system of equations

$$Y_{s,t} = \beta * DTT_{s,t} + \delta' X + \rho_t + \mu_s \quad (9)$$

where $Y_{s,t} = \left(FDI_{s,t}, Tax_{s,t}\right)'$ is a vector including the logarithm of the aggregate stock of foreign direct investment in country s and year t, $FDI_{s,t}$ and the logarithm of the aggregate corporate income tax revenue in country s and year t, $Tax_{s,t}$. The indicator vector variable $DTT_{s,t}$ equals one for country s if a treaty with Mauritius is effective in year t and zero otherwise. We include country-specific fixed effects $\mu_s = (\mu_{fdi,s}, \mu_{tax,s})$ to account for differences in FDI and tax revenue that exist independent from the tax treaty framework. Accordingly, the coefficient vector $\beta = (\beta_{fdi}, \beta_{tax})$ measures the relative change in FDI and the relative change in tax revenue, following treaty conclusion with Mauritius. These DD estimates of the treatment effect build on the premise that treated and control group would have moved in the same direction in the absence of an intervention; the common-trend assumption (see Angrist and Pischke, 2009). While there is no test to verify this assumption formally, we review whether the experimental groups followed similar trajectories before treaty-conclusion.

We include a broad set of control variables (X) that may affect FDI and tax revenues, drawing on the World Bank's World Development Indicators (WDI) and the IMF's WEO database. Specifically, we control for the magnitude of international trade, by including the log of imports and exports, market size and potential demand in source countries, including GDP and GDP per capita, GDP growth, and measures to control for the overall governance environment.[31] To account for resource prices, we include the price of oil, agricultural raw material, and a metal price index. Finally, we include the lagged value of FDI in the FDI equation. Due to the long timeframe analyzed in these equations, we are less concerned about bias introduced through the dynamic structure of one equation (Nickel, 1970). However, we further examine the issue as part of our robustness checks by substituting the stock of FDI with FDI inflows, so that the dynamic nature of the equation is removed.

Since disturbance terms in the FDI and tax revenue equation are potentially correlated, we estimate the system of equations in a seemingly unrelated regression (SUR) framework (see Zellner, 1962; Zellner and Huang, 1962; and Zellner, 1963). Specifically, we first estimate system (9) equation-by-equation to obtain estimates of the cross-equation correlation of residuals. Subsequently, we use this information to compute the system's covariance matrix and re-estimate the equations jointly in a generalized least squares approach. We thus improve estimation efficiency over the simpler OLS. We report Newey-West type standard

[31] Control of corruption and governance effectiveness.

errors (see e.g. Petersen, 2009) that allow for arbitrary heteroscedasticity and serial correlation within countries and across equations when not including a vector of time-fixed effects. In equations where we do include time-effects, we allow for heteroscedasticity and cross-sectional correlation when computing standard errors.

B. Estimation of Treaty-Shopping

DTTs have mechanical effect on revenue through the application of reduced tax rates on existing income flows. They may also trigger a rerouting of investment flows and other behavioral responses, such as an adjustment in the transfer pricing of internal MNE group transactions. We employ our conceptual framework to disentangle revenue losses due to income rerouting from revenue losses that would have also occurred in the absence of treaty-shopping. The intuition of our identification strategy can be illustrated in a two countries scenario, where both countries have the same average withholding tax rate on international income flows. While one country applies a single rate consistently across all investors, the second country applies a low rate on some investors and a high rate on others. If MNEs exploit such differential treatment, we expect tax revenue to be smaller in the second country. In turn, the lower tax burden reduces MNEs' cost of capital, thus increasing the optimal investment magnitude in the second country.

Correspondingly, we identify treaty shopping by measuring the systematic response of aggregate FDI and corporate income tax revenue to a country's variation in the treatment of cross border income flows, holding average tax rates constant. To identify the hypothetical cost parameter, we additionally assume (i.) that real investment potential is driven by a resident country's GDP,[32] and (ii.) that investor locations for SSA include OECD economies, China, India and 10 investment hubs.

Building on these assumptions, and the conceptual framework,[33] we estimate the following system of equations:

$$Y_{s,t} = f(z_{s,t}; \beta) + \delta' X_{s,t} + \rho_t + \mu_s + \varepsilon_{s,t}, \qquad (11.a)$$

where our two dependent variables remain the logarithm of the aggregate stock of foreign direct investment, FDI, and the logarithm of aggregate CIT revenue, Tax, in country s and year t. We account for the same set of explanatory variables, X, we used in our baseline regression results, a vector of time-specific fixed effects, ρ_t, and a vector of country-specific fixed effects μ_s. In addition, based on an approximation of equations 5 and 7, we include the vector-valued function:

[32] See Crivelli and others (2017) for a similar approach.

[33] Besides optimization behavior, core assumptions of the theory include that marginal treaty-shopping costs are proportional to the amount of rerouted income and that investments are financed with debt when made from low-tax jurisdictions.

$$f(\mathrm{z_{s,t}};\beta) = \begin{matrix} \beta_1 \frac{1}{1-t_s}\left[1 + \sum_i \bar{\alpha}_{t,i}\tau_{s,t,i} - \gamma\frac{1}{2}\frac{1}{N}\sum_i\left(\tau_{s,t,i} - \bar{\tau}_{s,t}\right)^2\right] \\ \beta_2\left[\sum_i \bar{\alpha}_{t,i}\eta_{s,t,i} - \gamma\frac{1}{N}\sum_i \eta_{s,t,i}\left(\tau_{s,t,i} - \bar{\tau}_{s,t}\right)\right] \end{matrix} \quad (11.b)$$

where $z_{s,t} = (t_s, \tau_{st}, \eta_{st}, \bar{\alpha}_{t,i})$ summarizes observables and $\beta = (\gamma, \beta_1, \beta_2)$ is a vector of parameters we seek to estimate. We use GDP weights to proxy for the real funding potential $\bar{\alpha}_{t,i}$ of investor country i in year t and use definitions (6.a) and (6.b) to compute the tax costs of income repatriation, $\tau_{s,t,i}$, and the domestic tax on international income flows, $\eta_{s,t,i}$, for all source countries s investor locations i in years t. Further we use statutory CIT rates to proxy for t_s.

Equation (11) captures our model's main prediction that variation in the treatment of cross-border income flows provides tax avoidance opportunities which exert a positive effect on FDI and a negative effect on tax revenue where MNEs exploit these opportunities. Accordingly, we expect that $-\beta_1\gamma > 0$ while $-\beta_2\gamma < 0$. Furthermore, we expect a negative coefficient β_1, reflecting that a higher average effective tax cost of income repatriation discourages investment. The coefficient β_2 is expected to be positive, indicating that a higher average effective tax rate on international income flows increases tax revenue.

As the unobservable treaty-shopping cost parameter impacts both equations, we estimate system (11) jointly in a GLS framework and account for its non-linear nature by iteratively minimizing the sum of squared residuals $\varepsilon_{s,t}$. To examine whether treaty-shopping costs are smaller for MNEs operating in countries that have concluded a DTT with an investment hub, we allow the unobservable cost parameter to vary in subsequent specifications.

C. Descriptive Statistics

We use aggregate (inward) FDI stock and tax revenue data for 41 countries in Sub-Saharan Africa between 1985 and 2015 as dependent variables. Our primary data source is the IMF's World Economic Outlook database, which in its extended version also covers detailed tax revenue data. Our primary independent variables draw on the IBFD database, which records information on DTTs globally, including on whether and in what year treaties became effective as well as the maximal withholding tax rates bilaterally agreed upon for different types of income payments. We construct a set of global CIT rates by combining the IMF's WoRLD database with information provided by KPMG and PwC.

Table 2 provides an overview on our explanatory and dependent variables. The variable Mauritius treaty is an indicator variable, taking the value of one if a tax treaty between a country in our sample and Mauritius is effective in a particular year. On average, 15 percent of observations in our sample are linked to an effective treaty with Mauritius. The variable Non-Mauritius treaty counts the number of effective treaties except with Mauritius. Similarly, the variables Hub-treaty and Non-hub treaty are count variables for the number of

effective treaties with either Mauritius or countries where the sum of FDI stocks exceeds 2 times the country's GDP in 2016.[34]

Observations on FDI are more complete than on corporate income tax revenues. Among countries that have concluded a treaty with Mauritius, the average logarithmic inward stock of FDI is at around 0.43, while it is only -0.65 among countries not having concluded such a treaty. Similarly, when not controlling for any other aspects, CIT revenues tend to be larger in countries that have concluded a treaty with Mauritius (-1.83) than those that have not (-2.57).

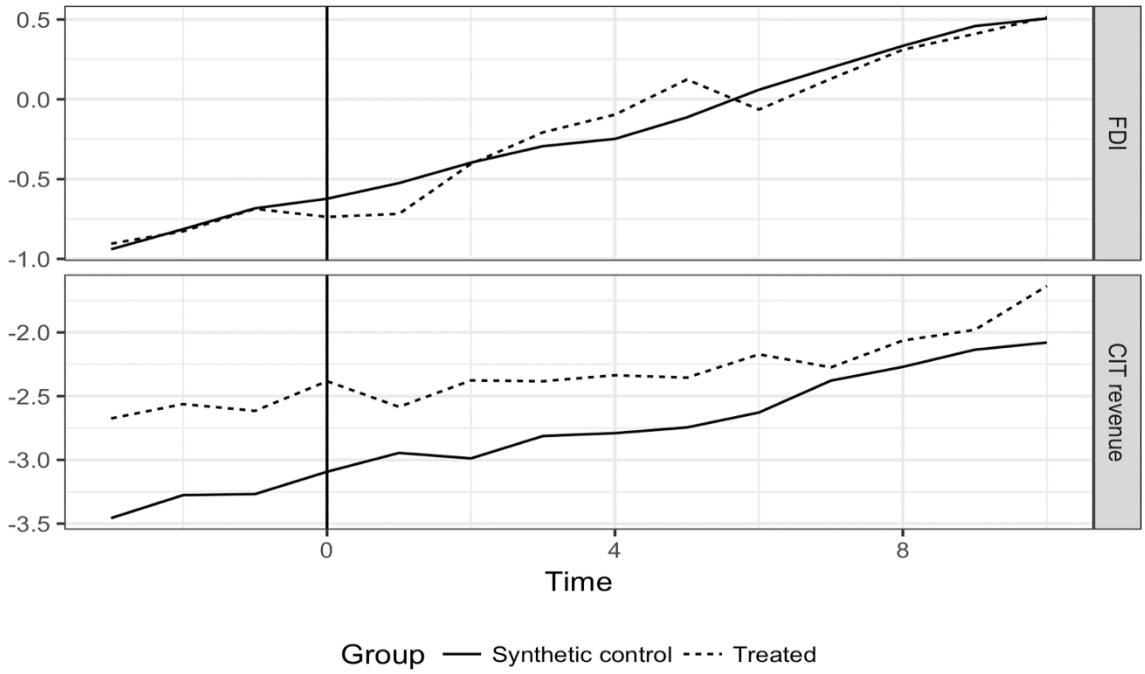
Table 2. Descriptive Statistics

Variable	Observations	Min	Mean	Max
Tax variables				
Mauritius treaty	1219	0	0.15	1
Non-mauritius treaty	1219	0	2.22	15
Hub-treaty	1219	0	0.27	3
Non-hub treaty	1219	0	2.1	13
CIT rate (in %)	1218	10	35.83	70
$\sum_i \bar{\alpha}_{t,i}\tau_{s,t,i}$	1219	-0.34	-0.08	0.11
$\sum_i \bar{\alpha}_{t,i}\eta_{s,t,i}$	1219	-0.59	-0.18	0.02
$\frac{1}{N}\sum_i \eta_{s,t,i}(\tau_{s,t,i} - \bar{\tau}_{s,t})$	1218	0	0	0.01
$\frac{1}{N}\sum_i(\tau_{s,t,i} - \bar{\tau}_{s,t})^2$	1218	0	0.01	0.02
Key control variables				
Log Imports	1219	-2.87	0.39	4.51
Log Exports	1219	-4.59	-0.03	4.63
Log Gdp	1219	-2.4	1.41	6.34
Gdp growth	1219	-0.96	0.06	0.78
Log Population	1214	12.68	15.73	19.04
Inflation	1209	-1.88	10.76	55.21
Exchange rate	1207	0.26	376.1	1662.69
Dependent variables				
Log Fdi	1196	-18.42	-0.48	4.55
Log Cit	632	-9.6	-2.43	2.88
Log Fdi with Mauritius DTT	182	-2.67	0.43	3.46
Log Cit with Mauritius DTT	119	-4.02	-1.83	0.14
Log Fdi without Mauritius DTT	1014	-18.42	-0.65	4.55
Log Cit without Mauritius DTT	513	-9.6	-2.57	2.88

Figure 2 summarizes FDI and CIT revenue for countries that concluded a DTT with Mauritius (Treated) and a synthetic control group which is chosen to resemble these countries before treaty-conclusion. Annex 2 provides details on our matching approach. The horizontal axis depicts years relative to the treaty with Mauritius becoming effective (Time 0). The graph suggests that both FDI and CIT have followed similar trends across these groups before treaties with Mauritius became effective, supporting the common-trend assumption. Despite selecting a control group which resembles treated countries, CIT revenue is noticeably smaller among control group members before the intervention, underscoring the importance to test treaty-effects in a difference in difference framework.

[34] Malta; Cyprus; Hong Kong SAR, China; Luxembourg; Singapore; Ireland; Switzerland; the Netherlands; Barbados; and the Seychelles.

Figure 2. Trends in FDI and CIT Revenue before the Intervention



Notes: Figure 2 presents average progression in the logarithmic stock of FDI and log CIT revenue for a group of countries that concluded a DTT with Mauritius (Treated) and a synthetic control group that has not concluded a treaty (Synthetic control). See Annex 2 for details on our matching approach. The horizontal axis depicts years relative to the treaty with Mauritius becoming effective (year 0).

V. RESULTS

A. Difference-in-Differences Based Results on Treaty Conclusion with Investment Hubs

Table 3 below summarizes our findings across four different SUR specifications that jointly estimate the determinants of FDI and of tax revenue, based on equation (9). We present heteroscedasticity- and serial correlation robust standard errors in parentheses. Results for the determinants of the FDI stock are summarized in the first four columns; columns 5 to 8 present the same for CIT revenue.

Overall, our estimations suggest that an additional tax treaty between source countries in SSA and Mauritius did not increase FDI. This finding is robust to expanding our explanatory variables. Notably, however, we do find a negative effect on CIT revenue which is also robust across our specifications. Estimated revenue losses are significant both economically and statistically and range between 15 and 25 percent of CIT revenue. The data does not allow differentiating between total CIT revenue and the share of revenue related to withholding taxation. However, the large magnitudes of captured effects suggest that the general decline in revenue is likely due to a combination of reduced WHT revenue and changed incentives for profit shifting. These observed effects exceed prior estimates of the revenue cost of aggressive tax evasion and avoidance and need to be interpreted with caution. Notably, our analysis identifies marginal effects of treaty-conclusion while actual effects

might be non-linear, in which case our estimates of total revenue losses are less precise

The first SUR model tests the partial effect of the binary indicator *Mauritius treaty*. The result indicates that CIT revenue decreased on average by around 25 percent following treaty conclusion with Mauritius. We include the number of other treaties a country has signed as an additional control variable in the second specification. While this variable does impact the total FDI stock positively, the investment effect of other treaties is measured imprecisely and not statistically significant. In contrast, we do find that the number of treaties a country has concluded impacts CIT revenue. Specifically, an additional treaty reduces tax collections by around 5 percent for the countries in our sample. While the point estimate of 18 percent suggests that revenue losses are more substantial for treaties concluded with Mauritius, the difference to other treaties is not statistically significant. In the third model, we include a vector of time-specific fixed effects. The negative effect of other treaties on CIT revenue is now estimated to be slightly larger at 5.5 percent, while the partial effect of concluding a treaty with Mauritius is slightly smaller, but still significant and sizeable at 15 percent. In the last column we drop insignificant explanatory variables to increase estimation precision. Our results remain unchanged.

All control variables in the four estimations follow theoretical predictions. Higher corporate income tax rates affect FDI negatively, as they reduce the net return to investments, while they increase tax revenue. Trade openness impacts positively on both FDI and revenue collection. A higher level of GDP is associated with less FDI, possibly due to a smaller return on investments in more developed countries. On the other hand, the positive effect of GDP in the revenue equation confirms that high income countries tend to collect more CIT revenue.

To test the importance of functional form assumptions, we re-estimate the effect of treaty conclusion with Mauritius using a range of matching estimators in Annex 2. Overall, estimated treatment effects closely follow the findings of our baseline regressions. Independent of the matching method used, the measured effects of treaty-conclusion on FDI are small and not statistically significant. In contrast, we find negative and significant effects on CIT revenue which resemble our baseline regression results.

Results for investment hubs We test the effect of treaty conclusion with investment hubs more generally by substituting the Mauritius indicator with the variable *Investment hub*, which counts the number of treaties that source countries have concluded with Mauritius; Malta; Cyprus; Hong Kong SAR, China; Luxembourg; Singapore; Ireland, Switzerland; the Netherlands; Barbados; and the Seychelles. Following our baseline estimations, we estimate the FDI and revenue equations jointly in an SUR framework. As before, the first four columns in Table 4 present the determinants of the FDI equation, the second four columns depict revenue effects.

Table 3. Baseline Results - The Effects of Signing Tax Treaties with Mauritius

SUR Estimation with GLS								
Dependent variable	Log (FDI)				Log(CIT revenue)			
Explanatory	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Mauritius treaty	-0.025	-0.034	-0.026	-0.035	-0.248**	-0.184 *	-0.155 *	-0.168*
	[0.031]	[0.032]	[0.032]	[0.038]	[0.098]	[0.101]	[0.092]	[0.096]
Non-Mauritius treaty		0.006	0.007	0.007		-0.048***	-0.055***	-0.043***
		[0.005]	[0.005]	[0.005]		[0.013]	[0.014]	[0.013]
CIT rate	-0.004 *	-0.003 *	-0.003		0.024***	0.019***	0.016***	0.019***
	[0.002]	[0.002]	[0.002]		[0.005]	[0.005]	[0.005]	[0.005]
Lag Log FDI	0.935***	0.933***	0.936***	0.936***				
	[0.029]	[0.030]	[0.030]	[0.026]				
Log Imports	0.129***	0.125***	0.129***	0.149***	-0.027	0.03	0.056	
	[0.040]	[0.040]	[0.039]	[0.045]	[0.151]	[0.154]	[0.145]	
Log Exports	0.005	0.005	-0.003		0.664***	0.631***	0.581***	0.628***
	[0.053]	[0.053]	[0.057]		[0.087]	[0.086]	[0.082]	[0.074]
Log Gdp	-0.134**	-0.131 *	-0.127 *	-0.134***	0.797***	0.782***	0.901***	0.894***
	[0.067]	[0.068]	[0.071]	[0.043]	[0.128]	[0.126]	[0.136]	[0.131]
Gdp Growth	0.148**	0.151**	0.161**	0.147**	0.034	0.045	0.048	
	[0.060]	[0.060]	[0.073]	[0.069]	[0.158]	[0.157]	[0.180]	
Log Population	0.236**	0.231**	0.287		0.653 *	0.812**	0.748	
	[0.102]	[0.101]	[0.195]		[0.348]	[0.324]	[0.490]	
Inflation	0.001	0.001	0.002		0.001	0.002	0.005**	
	[0.001]	[0.001]	[0.001]		[0.002]	[0.002]	[0.002]	
Exchange rate	0.008	0.009	0.011		-0.171***	-0.170***	-0.227***	-0.241***
	[0.014]	[0.014]	[0.013]		[0.059]	[0.058]	[0.055]	[0.054]
Control of corruption	0.087	0.089	0.1		0.138	0.165	0.18	
	[0.065]	[0.065]	[0.065]		[0.123]	[0.123]	[0.119]	
Government effectiveness	0.006	0.006	-0.012		-0.202	-0.262 *	-0.305**	-0.289**
	[0.054]	[0.054]	[0.053]		[0.141]	[0.136]	[0.130]	[0.117]
Oil price index	-0.002 *	-0.002 *			0	0		
	[0.001]	[0.001]			[0.001]	[0.001]		
Producer price index	-0.003	-0.003			-0.013***	-0.012**		
	[0.002]	[0.002]			[0.005]	[0.005]		
Metal price index	0.003**	0.003**			0.007**	0.007**		
	[0.002]	[0.002]			[0.003]	[0.003]		
Observations	1156	1156	1156	1190	614	614	614	617
Time FE	No	No	Yes	Yes	No	No	Yes	Yes

*, **, and *** indicate significance at the 10%, 5%, and 1% level. Heteroscedasticity and autocorrelation robust standard errors in square brackets.

While the point estimates reported in the previous section remain robust to extending the definition of investment hubs, the effects are now measured more precisely due to the increased variation in our main explanatory variable. In Model 1 and 2, we estimate the effect of concluding double tax treaties with investment hub and non-investment hub locations. We include a set of time-specific fixed effects in the second specification. Regardless of the approach, we do find sizeable negative effects of investment hub treaties on government revenues in source countries (-22 percent) but no additional effect on investment. The negative revenue effect for non-investment hub treaties is smaller (-3 percent) but also significant at the 5 percent and 1 percent levels. Notably, the estimation results indicate that revenue losses are significantly larger when the treaty partner country is an investment hub.

The estimation of a dynamic model with least squares introduces bias to all estimated coefficients (Nickel, 1970).[35] In the third specification, we thus examine whether our results are biased by the inclusion of lagged FDI values. To do so, we use FDI inflows as a share of GDP as our dependent variable but continue to use the same set of explanatory variables. The substitution of dependent variables does not affect our findings: DTTs do not impact investments significantly but reduce government revenue. The negative revenue effect remains considerably larger for investment hub treaties than for others.

In the fourth specification, we quantify the impact of DTTs on general income tax revenue, which includes personal income tax revenue and corporate income tax revenue. This measure is potentially less precise, as we expect the most direct effect on corporate taxes. However, by using a more aggregate revenue measure, we increase the number of observations by roughly 50 percent, from 614 to 918. Our results indicate that general income tax revenue decreases by 19.8 percent in response to each treaty concluded with an investment hub and by 0.9 percent in response to other treaties.

Table 4. Results for Investment Hubs and Robustness checks

SUR Estimation with GLS								
Dependent	Log (Investment)				Log (Tax Revenue)			
Explanatory	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Investment hub treaty	0.001	-0.001	-0.003	0.001	-0.237***	-0.217***	-0.257***	-0.198***
	[0.018]	[0.020]	[0.005]	[0.018]	[0.071]	[0.073]	[0.071]	[0.074]
Non-investment hub treaty	0.007	0.005	0.001	0.007	-0.026**	-0.035***	-0.027**	-0.009
	[0.005]	[0.005]	[0.001]	[0.005]	[0.011]	[0.011]	[0.012]	[0.008]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1190	1190	1158	1156	617	617	617	918

*, **, and *** indicate significance at the 10%, 5%, and 1% level. All specifications include the same set of explanatory variables as the last specification of the baseline results. Models 1 and 2 test for a broader definition of investment hubs, Model 3 investigates dynamic bias, and Model 4 quantifies effects on total income tax revenue. Heteroscedasticity and autocorrelation robust standard errors in square brackets.

[35] However, the bias vanishes with increasing time observations and our panel covers up to 25 years.

B. Estimation of Treaty-Shopping

Next, we analyze the importance of treaty-shopping in driving the observed effects, by estimating system (11) with non-linear least squares. We account for cross-equation correlation by joint estimation in a GLS framework. In all specifications, we include the same set of covariates as in our baseline specification. We alternate between a set of price indices and a vector of time-specific fixed effects to control for factors which affect all countries in a given year. For simplicity, we only report the core parameter estimates: the unobservable treaty-shopping costs parameter, γ, the marginal effect of taxation on investment, β_1, and the marginal effect of taxation on revenue, β_2.

Table 5 presents results. Throughout our specifications, the structural parameters have the expected sign, indicating consistency between the data and our conceptual framework: we find positive estimates of the treaty-shopping cost parameter, suggesting that countries with more variation in their withholding tax structure report a higher stock of FDI and less revenue. Moreover, we find that higher average tax costs of income repatriation impact FDI stocks negatively and that higher average tax rates on international income flows impact government revenues positively. While the estimated treaty-shopping parameter and the marginal effects of the average effective tax rate on international income flows is statistically significant in all specifications, the effect of repatriation taxes on the stock of FDI is measured less precisely and often not statistically significant at conventional levels.

Models 1 and 2 estimate the unobservable treaty-shopping cost parameter separately for the FDI and for the tax revenue equation. In both specifications, we do find that the parameter estimates follow our model's predictions, thus providing indirect evidence for treaty shopping activity. Variation in the tax costs of income repatriation increase FDI and reduce government tax collections. When including a vector of time-specific fixed effects in the second specification, both parameter estimates are statistically significant. As we are not able to reject the hypothesis that the structural cost parameter is identical in the two equations, we increase estimation precision by imposing this restriction in the third and fourth column. While results remain broadly unchanged, the coefficient measuring the effect of repatriation taxes on investment levels is measured more precisely and significant at the 10 percent level in Model 3. Estimation precision decreases when including vectors of time-specific fixed effects in Model 4.

We next examine whether countries are confronted with more income being rerouted following the conclusion of treaties with investment hubs. We expect that the measured cost parameter is smaller in countries that do not have a direct link with an investment hub than in countries that have concluded a DTT with an investment hub.[36] Accordingly, Models 5 and 6 estimate a separate treaty-shopping cost parameter for countries which signed a DTT with an investment hub and for those that did not. Irrespective of whether we use a set of price indices to control for common shocks (Model 5) or whether we use a vector of time-specific fixed effects (Model 6), we do find statistically significant evidence for this conjecture. In Model 5, the cost parameter is estimated to be 4.3 among countries that have not concluded a

[36] We expect this inequality as the parameter is inversely related to MNEs' overall costs of treaty shopping.

treaty with an investment hub and 16.8 (=4.3+12.5) among countries that have concluded such a treaty. The difference is significant at the 10 percent level. When including time-specific fixed effects, the parameter estimates barely change and the difference in treaty-shopping costs remains statistically significant.

Table 5. Estimation of Treaty-Shopping

Parameter estimate	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
γ_{fdi}	14.011**	23.088***				
	[6.141]	[6.562]				
γ_{tax}	3.87	5.103 *				
	[2.559]	[2.986]				
$\gamma_{combined}$			4.198 *	5.355 *		
			[2.451]	[2.954]		
γ_{nT}					4.301 *	5.605 *
					[2.470]	[2.990]
$\gamma_{yT} - \gamma_{nT}$					12.529 *	11.214 *
					[6.780]	[5.768]
β_1	-0.779	-0.616	-0.472 *	-0.245	-0.461 *	-0.247
	[0.476]	[0.421]	[0.268]	[0.230]	[0.277]	[0.215]
β_2	9.351**	10.445***	9.429**	10.487***	9.211**	10.216***
	[3.925]	[2.726]	[3.940]	[2.727]	[3.983]	[2.640]
Time fixed effects	No	Yes	No	Yes	No	Yes
Firm-specific fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

*, **, and *** indicate significance at the 10%, 5%, and 1% level. All specifications include the same set of explanatory variables as the last specification of the baseline results. Models 1 and 2 estimate the parameter γ separately for the FDI and tax revenue equation, Model 3 and 4 impose a common parameter, Model 5 and 6 allow the parameter to differ depending on whether the observational unit has concluded a treaty with an investment hub (yT) or not (nT). Heteroscedasticity and autocorrelation robust standard errors in square brackets.

Interpretation of Structural Results

We combine the structural parameter estimates with our theoretical model to assess the extent and consequences of treaty-shopping. We present three summary statistics: the share of total investments in source countries which is financed indirectly to reduce MNE's tax bill (diversions), resulting revenue losses and the implied investment gains.

The share of diverted income in country s is given by $\sum_i |\alpha_{s,i} - \overline{\alpha_{s,i}}|$, where $\alpha_{s,i}$ denotes the share of reported funding from location i and $\overline{\alpha_{s,i}}$ denotes real funding share in this location. Using equation (4), we can rewrite this sum in terms of observables and the estimated treaty-shopping cost parameter:

$$Diversion_{s,t} \equiv \gamma \frac{1}{N} \sum_i \left| \tau_{s,t,i} - \bar{\tau}_{s,t} \right| .$$

The variable diversion takes on values between zero and one. A value of zero implies that all investments are directly financed from places where the real funding potential is located. In contrast, a value of one implies that all income flows are rerouted to exploit a preferential tax treatment in locations other than where the real funding potential lies.

Treaty-shopping opportunities reduce government revenue and, in turn, increase FDI. To quantify investment gains and revenue losses resulting from treaty-shopping we use equation (11) and compute

$$Investment\ Gain_{s,t} \equiv -\beta_1 \gamma \frac{1}{2} \frac{1}{N} \sum_i \left(\tau_{s,t,i} - \bar{\tau}_{s,t}\right)^2 ,$$

$$Revenue\ Loss_{s,t} \equiv \beta_2 \gamma \frac{1}{N} \sum_i \eta_{s,t,i} \left(\tau_{s,t,i} - \bar{\tau}_{s,t}\right) ,$$

Table 6 summarizes the results. We first assess effects of treaty-shopping with the unobservable cost parameter being constant across all SSA economies (see column 1). Using the structural parameters of Model 3 in Table 5, the simulation suggests that some 36 percent of all income flows are rerouted to reduce tax payments, resulting in tax revenue losses of around 5 percent of CIT revenue and increased investments of around 1 percent of FDI across SSA. Next, we allow treaty-shopping costs to differ depending on whether countries have concluded a DTT with an investment hub in columns 2 and 3. Using the parameter estimates of Model 5 in Table 5, we find considerably more income being rerouted in the former group (around 90 percent) than in the latter (30 percent). Because of treaty shopping, countries with an investment hub treaty risk losing around 15 percent of CIT revenue, while the reduced cost of capital may attract around 3 percent more FDI. Note that revenue losses due to treaty-shopping are around 10 percent larger among countries that have concluded a treaty with an investment hub. Contrasting this effect with the estimates presented in Section 5.A suggests that rerouted income accounts for around half of total revenue losses linked to DTTs with investment hubs.[37]

These estimated magnitudes appear large compared to the existing literature on international tax avoidance. For instance, the OECD (2015) assesses revenue losses linked to base erosion and profit shifting between 4-10 percent globally. However, country-specific results can be much larger,[38] and empirical work has thus far primarily focused on arbitrage opportunities between countries' statutory CIT rates. In contrast, we examine revenue losses related to treaty-shopping. Our findings are in line with recent evidence on the importance of offshore centers globally (see Damgaard and others 2018), indicating that more than half of all outward FDI flows pass through foreign shell companies. For SSA, we show that DTTs critically impact the extent of these diversions."

[37] The DD estimates indicate that total revenue losses due to investment hub treaties are around 20 percent of CIT revenue.

[38] Beer and others (2018) estimate that MNEs' tax avoidance might reduce CIT revenue in the US by 17 percent; Clausing (2016) puts this number between 19 and 30 percent.

Table 6. Simulation Results – Effect of Treaty-Shopping			
Column	(1)	(2)	(3)
Subsample	All countries	DTT with investment hub	No DTT with investment hub
Diversion (in%)	36.00	92.71	30.20
FDI Gain (in%)	0.96	2.90	0.76
CIT Revenue Loss (in%)	5.13	14.48	4.17

Notes: Based on the conceptual framework and estimation results of Table 5, this Table presents simulation results in percent, differentiating between three country groups.

VI. CONCLUSION

We analyze the cost and benefits of concluding Double Tax Treaties (DTTs) with investment hubs on FDI and government revenues for a sample of African economies. Our results imply that these economies are likely facing a substantial reduction in revenue when signing a treaty with investment hubs. At the same time, we find no evidence substantiating expectations of a corresponding benefit in terms of additional investment generation. The value of tax treaties to attract investors thus remains uncertain at best, especially where investments tend to be aimed at a country's natural resources or specific market.

Further analysis is needed to assess the large magnitudes of treaty related base erosion and profit shifting in Sub-Saharan Africa identified in our analysis. While micro-level firm level information on MNEs remains very limited in commercial databases for many developing countries, including most economies in Sub-Saharan Africa (Platform for Collaboration on Tax 2017), future research drawing on micro level administrative information, including Country-by-Country reports of large MNEs, could potentially help fill this gap

To assess the robustness of our findings and evaluate their potential underlying drivers, we examine the role of treaty shopping in driving nominal investment flows and provide initial empirical evidence on the importance of treaty-shopping in the region. Our analysis suggests that the conclusion of DTTs with investment hubs increases the rerouting of income flows. These results confirm that avoidance opportunities have a major impact on the overall cost-benefit ratio of tax treaties. Ongoing policy reforms and international initiatives to include or improve anti-abuse provisions in tax treaties, including those signed with Mauritius, may help address this challenge.

The findings of this paper are relevant for tax policy makers in Sub-Saharan Africa when balancing revenue and investment promotion objectives and for wider discussions on the international tax architecture. Tax treaties can be beneficial where their benefits outweigh costs. Our paper raises further doubts on this being the case for treaties concluded with investment hubs.

References

ActionAid (2016)/Hearson, M. The ActionAid Tax Treaties Dataset

Alstadsaeter, A., N. Johannensen, and G. Zuman. (2017). Who Owns the Wealth in Tax Havens? Macro Evidence and Implications for Global Inequality, NBER Working Paper 23805.

Angrist, J. and J. Pischke. (2009). Mostly Harmless Econometrics: An Empiricist's Companion, Princeton: Princeton University Press.

Balabushko, O., S. Beer, J. Loeprick, F. Pinto Vallada. (2017). The Direct and Indirect Costs of Tax Treaty Policy: Evidence from Ukraine, Policy Research Working Paper Series 7982, The World Bank.

Barthel, F., Busse and E. Neumeyer. (2010). The Impact of Double Taxation Treaties on Foreign Direct Investment: Evidence from a large Dyadic Panel Data. Contemporary Economic Policy Volume 28, Issue 3, pages 366–377, July 2010.

_____ and E. Neumeyer. (2012). Competing for Scarce Foreign Capital: Spatial Dependence in the Diffusion of Double Taxation Treaties, International Studies Quarterly, 56: 645–660.

Beer, S, R. De Mooij, and L. Liu. (2018). International corporate tax avoidance: a review of the channels, effect size, and blind spots.

Bergstrand, J. H. and P. Egger. (2007). A Knowledge-and-Physical-Capital Model of International Trade Flows, Foreign Direct Investment, and Multinational Enterprises, Journal of International Economics, 73(2): 278-308.

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 2015 123:6, 1311-1355.

Braun, J. and M. Zagler. (2014). An Economic Perspective on Double Tax Treaties with(in) Developing Countries. World Tax Journal, Vol 6, No 3.

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

Carr, D., J.R. Markusen, and K.E. Maskus. (2001), Estimating the Knowledge–Capital Model of the Multinational Enterprise, American Economic Review, Vol. 91(3): 693–708.

Clausing, K. 2016. The Effect of Profit Shifting on the Corporate Tax Base in the United States and Beyond, *National Tax Journal*, 69(4), pages 905-934.

Crivelli, E., R. De Mooij, and M. Keen. (2016). Base Erosion, Profit Shifting and Developing Countries. FinanzArchiv: Public Finance Analysis, 72(3).

Damgaard, J., and T. Elkjaer. (2017). The Global FDI Network: Searching for Ultimate Investors, IMF Working Paper 17/258.

_____, T. Elkjaer, and N. Johannesen (2018), Piercing the Veil. Finance and Development, June 2018, pp. 50-53.

De Mooij, Ruud and Li Liu. 2018. At a Cost: The Real Effects of Transfer Pricing Regulations, IMF Working Paper 18/69

EDB. 2018. The Mauritius IFC: Driving Quality Investments in Africa. http://edbmauritius.org/files/Rebuttal%20ICIJ_MauritiusEDB.pdf

Egger, P., M. Larch., M. Pfaffermayr, and H. Winner. (2006). The impact of endogenous tax treaties on foreign direct investment: theory and evidence, Canadian Journal of Economics, Canadian Economics Association, vol. 39(3), pages 901-931, August.

Gravelle, J. G. (2013). Tax Havens: International Tax Avoidance and Evasion, Congressional Research Service Report for Congress. https://fas.org/sgp/crs/misc/R40623.pdf

Grubert, H. and Slemrod, J., (1998), The Effect of Taxes On Investment And Income Shifting To Puerto Rico, The Review of Economics and Statistics, 80, issue 3, p. 365-373.

Head, K. and J. Ries. (2008). FDI as an Outcome of the Market for Corporate Control: Theory and Evidence, Journal of International Economics, Vol. 74(1): 2-20.

Hearson, M. (2017). What makes Countries Negotiate away their Corporate Tax Base? WIDER working paper, 2017/122. United Nations University World Institute for Development Economics Research, Helsinki, Finland.

_____. (2015). Tax treaties in sub-Saharan Africa: A Critical Review, Tax Justice Network - Africa, Nairobi, Kenya.

ICIJ. (2017). Tax Haven Mauritius' Rise Comes at the Rest of Africa's Expense. November 7th. https://www.icij.org/investigations/paradise-papers/tax-haven-mauritius-africa/

Imai, K., In Song Kim, and Erik Wang (2018). Matching Methods for Causal Inference with Time-Series Cross-Section Data. Working Paper.

Janský, P., and Šedivý, M, (2018). Estimating the Revenue Costs of Tax Treaties in Developing Countries. IES Working Paper. Charles University. Forthcoming.

Lejour (2014). The Foreign Investment Effects of Tax Treaties'. CPB Netherlands Bureau for Economic Analysis Discussion Paper 265. The Hague: Netherlands Bureau for Economic Analysis.

McGauran, K. (2013). Should the Netherlands Sign Tax Treaties with Developing Countries? Somo. https://www.somo.nl/should-the-netherlands-sign-tax-treaties-with-developing-countries/

Neumayer, E. (2007). Do Double Taxation Treaties Increase Foreign Direct Investment to Developing Countries?, Journal of Development Studies, Taylor & Francis Journals, vol. 43(8), pages 1501-1519.

OECD. 2015. 2015 Final Reports, OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris

Petersen, MA. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, The Review of Financial Studies, 22(1), 435–480.

Petkova, K., A. Stasio, and M. Zagler. (2018). On the Relevance of Double Tax Treaties. WU International Taxation Research Paper Series, No. 2018-05.

Platform for Collaboration on Tax. 2017. A Toolkit for Addressing Difficulties in Accessing Comparables Data for Transfer Pricing Analyses. http://documents.worldbank.org/curated/en/447901498066167863/The-platform-for-collaboration-on-tax-a-toolkit-for-addressing-difficulties-in-accessing-comparables-data-for-transfer-pricing-analyses

Rixen, T., and P. Schwarz. (2009). Bargaining over the Avoidance of Double Taxation: Evidence from German Tax Treaties. FinanzArchiv: Public Finance Analysis, 65(4): 442–71.

Rosenbaum, P. R., and D. B. Rubin. (1983). Assessing Sensitivity to An Unobserved Binary Covariate in An Observational Study with Binary Outcome, Journal of the Royal Statistical Society, Series B, Methodological 45: 212-218.

Rubin, D. (2006). Matched Sampling for Causal Effects, Cambridge: Cambridge University Press.

Stuart, Elizabeth A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward, Statistical Science 25 (1): 1-21.

Torslov, T.R,, L.S. Wier, and G. Zucman. (2018). The Missing Profits of Nations, NBER Working Paper 24701.

UNCTAD. (2015). World Investment Report, Reforming International Investment Governance.

Van't Riet, M. and A. Lejour. 2017. Optimal tax routing: Network analysis of FDI diversion. International Tax and Public Finance (forthcoming), published online 18.04.2018.

Weichenrieder, A.J. and J.M. Mintz. (2008). What Determines the use of Holding Companies and Ownership Chains? Oxford University Centre For Business Taxation WP 08/03.

Weyzig, F. (2013). Tax Treaty Shopping: Structural Determinants of Foreign Direct Investment Routed through the Netherlands. International Tax and Public Finance, 20(6): 910–37.

World Bank (2015), Options for Low Income Countries' Effective and Efficient use of Tax Incentives for Investment: A Report to the G-20 Development Working Group by the IMF, OECD, UN and World Bank. Platform for Collaboration on Tax. Washington, D.C.: World Bank Group.

_____. 2017. Foreign Investor Perspectives and Policy Implications. Washington, D.C. World Bank Group.

Zellner. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, Journal of the American Statistical AssociationVol. 57, No. 298, pp. 348-368.

_____. (1963). Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results, Journal of the American Statistical Association, Vol. 58, No. 304, pp. 997-992.

_____ and Huang. (1962). Further Properties of Efficient Estimators for Seemingly Unrelated Regression Equations, International Economic Review, Vol. 3, No. 3, pp. 300-313.

Zucman, G. (2015), The Hidden Wealth of Nations – The Scourge of Tax Havens, Chicago.

ANNEX

Annex 1. Revenue-Maximizing Withholding Tax Rate

The first-order condition for an optimal withholding tax rate ω_i^{pas} vis-à-vis a low tax jurisdiction such that $\omega_i^{pas} \leq t_i^h$ is

$$\{t_s f'(K) + r[E(\eta_i) - \gamma Cov(\tau_i, \eta_i)]\}\frac{\partial K}{\partial \omega_i^{pas}} + \frac{\partial T}{\partial \omega_i^{pas}} = 0$$

where

$$\frac{\partial K}{\partial \omega_i^{pas}} = \frac{r}{1-\tau_s}\frac{1}{f''(K)}\left[\overline{\beta_i} - \gamma\frac{1}{N}(\tau_i - \bar{\tau})\right] \leq 0$$

$$\frac{\partial T}{\partial \omega_i^{pas}} = rK\left[\overline{\beta_i} - \gamma\frac{1}{N}\left[\tau_i - \bar{\tau} + \eta_i - \frac{1}{N}\sum_i \eta_i\right]\right] \geq 0$$

Note that for all locations i where $\omega_i^{pas} < t_i$, increasing the withholding tax rate vis-à-vis this location does not impact investment while it does increase revenue. Hence, an optimal withholding policy satisfies $t_i \leq \omega_i^{pas}$ for all i such that $t_i \leq t_i^H$, implying that $\tau_i = \eta_i$ for all i. So, $\frac{\partial T}{\partial \omega_i^{pas}} = rK\left[\overline{\beta_i} - \gamma\frac{2}{N}[\tau_i - \bar{\tau}]\right]$. Using this condition, and multiplying the necessary condition by $f''(K)$, we obtain

$$FOC: \; -A\frac{\gamma}{N}(\tau_i - \bar{\tau}) + B\beta_i = 0$$

where

$$A = \left(\frac{r}{1-t_s}\right)^2\left[t_s + E(\tau) - \gamma\frac{(2-t_s)}{2}Var(\tau)\right] + 2rKf''(K), \quad B = A - rKf''(K)$$

Concavity of the production function implies that $B \geq A$. We seek to define the properties of an interior solution where a DTT reduces cross-border taxation relative to other countries. Formally, we look for a solution satisfying $\tau_i^* < \bar{\tau}$. For an interior solution to exist in this region, we need $A < 0$ and $B \geq 0$, which holds if the production function is sufficiently convex and the costs of treaty-shopping are not close to zero.

The rate implicitly defined by FOC is a maximum if

$$SOC: \; -A\left(1 - \frac{1}{N}\right)\frac{\gamma}{N} - A'_\omega\frac{\gamma}{N}(\tau_i - \bar{\tau}) + B'_\omega\,\beta_i \geq 0$$

Note that the inequality is reversed due to the multiplication by f''. This condition holds if $\frac{f''(K)+Kf'''(K)}{f''(K)} > -\frac{1}{2}$, since then $B'_\omega > A'_\omega$ and

$$A'_\omega = \left(\frac{r}{1-t_s}\right)^2\left[\beta_i - \gamma\frac{(2-t_s)}{N}(\tau_i - \bar{\tau})\right] + 2r[f''(K) + Kf'''(K)]\frac{\partial K}{\partial \omega_i^{pas}} \geq 0$$

The optimal withholding tax rate is implicitly defined by $\tau_i = \bar{\tau} + \frac{B}{A}\frac{N}{\gamma}\beta_i$. Implicit differentiation gives

$$\frac{\partial \tau_i^*}{\partial \gamma} = \left[1 - \frac{1}{N} - \frac{\partial \frac{B}{A}}{\partial \omega_i^{pas}}\frac{N}{\gamma}\beta_i\right]^{-1}\left[\frac{\partial \frac{B}{A}}{\partial \gamma} - \frac{B}{A}\frac{1}{\gamma}\right]\frac{N}{\gamma}\beta_i > 0$$

Since $B'_\omega > A'_\omega > 0$ and $A < 0 < B$, it follows that $B'A - A'B < 0$, and thus $-\frac{\partial \frac{B}{A}}{\partial \omega_i^{pas}} > 0$. Further,

$$A'_\gamma = \left[-\left(\frac{r}{1-t_s}\right)^2 \frac{(2-t_s)}{2} - r[f''(K) + Kf'''(K)]\frac{r}{1-\tau_s}\frac{1}{f''(K)}\right] Var(\tau) < 0$$

And $B'_\gamma < 0$. Therefore, $B'_\gamma A - A'_\gamma B > 0$ implying that $\frac{\partial \frac{B}{A}}{\partial \gamma} > 0$.

Annex 2. Matching-Based Estimation

To test the importance of functional form assumptions, we re-estimate the effect of treaty conclusion with Mauritius using matching estimators (see Rubin, 2006; Stuart 2010). Intuitively, for each country that signed a treaty with Mauritius, we construct a separate control group that resembles the "treated" country in key dimensions, but where no treaty was concluded. Subsequently, we contrast the development of FDI and tax revenue between these groups to quantify the treatment effect non-parametrically.[39]

Following Imai, Kim, and Wang (2018), we identify the treatment effect of concluding a treaty with Mauritius F periods after the treatment using

$$\beta^F = \frac{1}{\sum_{i=1}^{N}\sum_{t=L+1}^{T-F} D_{it}} \sum_{s=1}^{N}\sum_{t=L+1}^{T-F} D_{it}\left[\left(Y_{s,t+F} - Y_{s,t-1}\right) - \sum_{j \in M_{itF}} \omega_{ijt}\left(Y_{j,t+F} - Y_{j,t-1}\right)\right] \quad (10)$$

where $D_{it} = 1$ if country i signed a DTT with Mauritius in period t and zero otherwise, Y is the outcome variable (either log of FDI or log of CIT revenue), M_{itF} is the set of potential control countries and ω_{ijt} are weights that capture similarity between country i and country j.

[39] Matching and regression-based estimators both presume that all confounding factors are observable or constant across time. While matching estimators relax parametric assumptions, thus potentially increasing the validity of causal inference (Ho and others, 2007), they require that the counterfactual outcome is a convex combination of observed outcomes among the control group. In contrast, regression-based analyses are less restrictive in this regard, allowing the extrapolation of counterfactuals. Finally, while the regression-based approach presumes a common trend between treated and controls, our matching-based approach allows trends to differ between countries. Consequently, results may differ.

The estimator depicted in equation (10) identifies the treatment effect as the difference between the average observed change in the outcome variable, F years after treaty conclusion with Mauritius (the first term in square brackets) and what would have happened to treated countries in the absence of concluding the treaty. This counterfactual outcome is estimated using a weighted average among control group members (the second term in square brackets). The time dimension in panel data requires a careful definition of the set of potential controls, since countries which conclude a treaty at a later point in time can be used as controls earlier:

$$M_{itF} = \{j: j \neq i, D_{jt'} = 0 \; for \;\; t' \leq t + F\} \tag{11}$$

The control set for country i, which was treated in year t, thus comprises all countries except country i which were not treated until F periods after year t. Figure A.1 illustrates how treated countries increase over time.

To assign weights and quantify similarity between treated and control observations, we use the synthetic control approach (Abadie, Diamond, Hainmueller, 2010) and pairwise Mahalanobis distances.[40] All weights are chosen with a view to achieve a balanced distribution of observable variables across treated and control group, so that the treatment dummy is observationally independent from other confounding factors. While the synthetic control approach maximizes covariate balance, an optimal balance may require assigning a weight of unity to one control observation and discarding all other observations. To test whether the implied loss of information affects results, we draw on pairwise Mahalanobis distances and construct two additional sets of weights using the 8 (20) most similar countries in the control group. We assess similarity using the vector of control variables included in the regression-based analysis as well as lagged values of the dependent variable and lagged values of the change in the dependent variable.

Note that regression-based identification implicitly presumes an instantaneous and constant treatment effect while the matching-based estimator depicted in equation (10) predicts the treatment effect for a specified number of periods after the treatment. To allow comparability of results, we average the dynamic effects over time using

$$\beta^{\bar{F}} = \frac{1}{F}\sum_{f=1}^{F} \beta^{f} \tag{12}$$

Following Imai, Kim, and Wang (2018), we compute standard errors using a block-bootstrap procedure with 500 iterations.

[40] A popular approach to quantify similarity and assign weights is the difference in the predicted likelihood of receiving treatment, the propensity score (Rosenbaum and Rubin, 1990). In panel settings, where the set of control units changes over time, propensity score matching requires estimation of a separate binary outcome regression for each period. This strategy is infeasible in our context, as in several years only one country concluded a treaty with Mauritius.

Matching-Based Estimation Results

We test the importance of parametric assumptions by re-estimating the treatment effect with a range of matching estimators.

Specifically, we first create control groups and compute weights that reflect similarity based on two different sets of control variables. One set of controls, the *small* set, includes only the lagged values of dependent variable as well as their change. The *large* set adds all explanatory variables that were used in the baseline regression results. Using equations (8) and (9), we then calculate the average treatment effect over three different time horizons, F=(3,5,10).

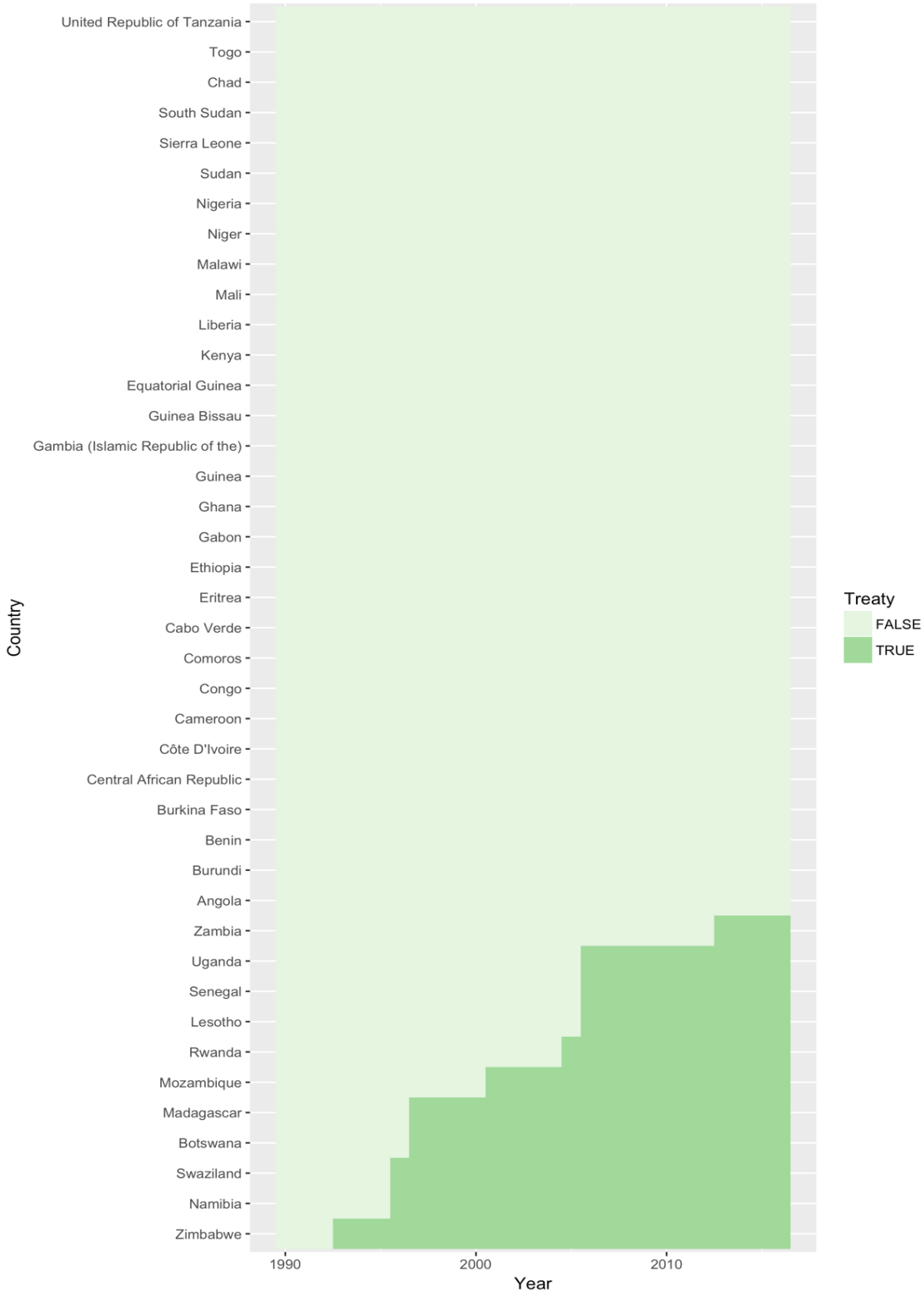
Figure 2 in the main text plots average values of our dependent variables for the treatment and the synthetic control. The figure suggests that the synthetic control groups followed a similar trend as the treatment group, both in terms of CIT revenue and FDI, before the intervention.

Table 3 presents matching estimator results. The top panel depicts matching estimators which are conditional on the small set of controls; the bottom panel presents estimators which are conditional on the large set of controls. The columns and rows differentiate between matching methods and time horizons.

Overall, estimated treatment effects closely follow the findings of our baseline regression framework. Independent of the matching method used, the measured effects of treaty-conclusion on FDI are small and not statistically significant. In contrast, we find negative and significant effects on CIT revenue. When using the small set of controls, the synthetic control approach indicates that treaty-conclusion reduces CIT revenue by an average of 25% within four years of signing a treaty (F+1); and by an average of 18% within 6 years. When using the larger set of controls, the estimated treatment effect increases to -32% (within 4 years) and to -25% (within 6 years). Mahalanobis-distance based matching implies larger effects.

Note that the number of potential control countries decreases as the evaluation horizon F increases. It is thus not surprising to find more precisely measured treatment effects over the shorter time horizon. In contrast, we do not find statistically significant results over the longer evaluation horizon. However, it is reassuring that estimated point estimates remain almost identical to the regression based identification.

Figure A.1



Notes: Figure presents when a DTT with Mauritius became effective in our baseline sample.

Table 7. Matched DD Estimators

Dependent variable		Log(Fdi)			Log(Cit Revenue)		
Matching method		M1	M2	Matching method	M1	M2	Matching method
Controls: small set				Controls: small set			Controls: small set
Time horizon	F=3	-0.052	-0.052	Time horizon F=3	-0.052	-0.052	Time horizon F=3
		[0.129)	[0.123)		[0.129)	[0.123)	
	F=5	-0.002	-0.01	F=5	-0.002	-0.01	F=5
		[0.156)	[0.137)		[0.156)	[0.137)	
Average for small set		-0.027	-0.031	Average for small set	-0.027	-0.031	Average for small set
Controls: large set				Controls: large set			Controls: large set
Time horizon	F=3	-0.098	-0.062	Time horizon F=3	-0.098	-0.062	Time horizon F=3
		[0.138)	[0.128)		[0.138)	[0.128)	
	F=5	-0.073	-0.04	F=5	-0.073	-0.04	F=5
		[0.162)	[0.153)		[0.162)	[0.153)	
Average for large set		-0.0855	-0.051	Average for large set	-0.0855	-0.051	Average for large set

Notes: Table presents treatment effects on log of FDI and log of CIT revenue based on DD matching. Square brackets give standard errors based on block-bootstrapping. Columns labeled M1 and M2 present results where matching is based on the Mahalanobis distance, including the nearest 8 (M1) or 20 (M2) countries. S presents matching results, using the synthetic control approach. Rows differentiate between the considered horizon (F). Top panel matches based on a small set of controls; bottom panel matches based on larger set of controls. See text for details.