DEVELOPING COUNTRY SECOND-MOVER ADVANTAGE IN COMPETITION OVER STANDARDS AND TAXES
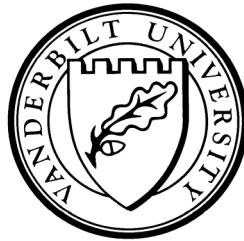
by

Valeska Groenert, Myrna Wooders, and Ben Zissimos

Developing Country Second-Mover Advantage in Competition over Environmental Standards and Taxes[1]

Valeska Groenert[2]
Vanderbilt University

Myrna Wooders[3]
Vanderbilt University

Ben Zissimos[4]
Vanderbilt University

First draft: October 2008
Revised: July 2009

ABSTRACT: We show that, in competition between a developed country and a developing country over environmental standards and taxes, the developing country may have a 'second-mover advantage.' In our model, firms do not unanimously prefer lower environmental-standard levels. We introduce this feature to an otherwise familiar model of fiscal competition. Three distinct outcomes can be characterized by varying the cost to firms of 'standard mismatch': (1) the outcome may be efficient; (2) the developing country may be a 'pollution haven,' where some firms escape excessively high environmental standards in the developed country; (3) environmental standards may be set excessively high.

KEYWORDS: Environmental standards, fiscal competition, second-mover advantage, tax competition.
JEL CLASSIFICATION NUMBERS: H2; H3; Q2

1 Introduction

The recent integration of countries in Eastern Europe to the European Union (EU) has provoked renewed concern about the aggressive competition by new members for firms and other mobile factors. For example, although EU accession requirements demand moves towards harmonization of environmental standards and some measures have made it onto statute books, there appears to be widespread skepticism about the actual implementation of such measures. Citing the incentive not to raise standards in order to attract firms, Post (2002) states that 'there is a "deception gap" between what is said on paper and what is done in practice' with regard to environmental policy.[5] To investigate this concern, our paper develops a model of international competition over environmental standards (ESs) and taxes. Firms who locate in a country are required to pay a tax that is used, at least in part, to enforce the ES in that country. The main purpose of this paper is to show that, through competition in ESs and taxes, a developing/transition country may indeed have a 'second-mover advantage' over a developed country in attracting firms and extracting rents. While this concern has circulated in policy discussions for some time now, to our knowledge it has not been studied formally before in the literature on fiscal competition.

This issue has been raised particularly with respect to the more economically successful 'transition countries' from the former Soviet Union as well as, to a lesser extent, the 'emerging market' developing countries in Asia and the Middle East. The so-called 'Visegrad countries' of the Czech Republic, Hungary, Poland and Slovakia (V4 for short) exemplify the developing and transition countries that we have in mind. These countries are in the midsts of comprehensive governmental reforms, and arguably their governments have a greater degree of flexibility and fewer constitutional and institutional constraints than the long-established democracies in the core of Europe. A greater degree of flexibility in policy-making has also been observed in dictatorships and young democracies further afield. For brevity, throughout the paper we will use 'developing country' as a catch-all term for all such countries.[6]

---

[5] Andanova (2003) provides further details of environmental policy in Eastern Europe.

[6] It should be understood that we are excluding from consideration a significant group of transition countries and less developed countries whose economic performances remain poor, not least because their policy-making processes are bogged down in a quagmire of distributional and special-interest concerns. The World Bank's (1996) World Development Report focuses specifically on a comparison in economic performance of the transition economies, grouping the twenty-six countries by numbers 1-4, with the top-performing V4 countries in group 1, etc. Specific details substantiating the distinction we make between governments in transition countries are provided in Chapter 7 of World Bank (1996), which focuses on government and policy formation; see especially pages 113-115.

We capture this greater flexibility in policy-making by the developing country, as a result of which it can respond (say within the period of a parliament) to policies adopted by the developed country, through our specification of timing in the game of policy formation. In our two country model, we will assume that the developed country sets its standard and tax first, followed by the developing country.

In contrast to the past literature (summarized below), we focus on a situation where the tax base is not universally repelled by, nor universally attracted to, ESs. Consider the familiar textbook example where an upstream firm pollutes a river, harming the profits of a fisherman downstream. An ES that requires the firm to reduce its pollution will improve the profits of the fisherman while reducing the profits of the firm. We extend this example to consider a general setting where one group of firms exerts a negative environmental externality on another group, and where the negative effect can be moderated through an ES. Formally, the set of firms can be ranked according to how the ES affects their profitability, where the effect on a particular firm can be positive or negative depending on the level of the standard. Thus each firm has an ideal ES level and this ideal level varies across firms. We introduce this feature to an otherwise familiar model of fiscal competition. It is this feature of the model that gives rise to the second-mover advantage in policy-setting that we identify.[7]

As the discussion so far suggests, we model competition for mobile firms as a sequential game between governments who choose standards and taxes. Due to monitoring costs, the higher the standard set by a country the more costly it is to implement. Following a common hypothesis in the literature (due to Niskanen 1977), national governments are run by bureaucrats who seek to maximize their budgets (tax revenue minus the cost of implementing the standard). There is a continuum of firms (while consumers are not explicitly considered).

---

[7]Broadly, the prior literature on interjurisdictional competition over ESs and taxes can be categorized into two areas. The first area, following Tiebout (1956), focuses on situations where competition among independent governments is like competition among firms and enhances efficiency. Here the 'Tiebout assumption' is that all firms benefit to differing degrees from a clean environment and sort themselves efficiently into jurisdictions each of which enforces an ES that is appropriate for its members. The second area concerns the presence of a policy-failure that allows or induces governments to set taxes on mobile capital, as in the literature on fiscal federalism and 'standard tax competition' associated with Oates (1972), Wilson (1986) and Zodrow and Mieszkowski (1986). Capital is indifferent to the imposition of an ES, but is repelled if burdened with having to pay for the ES. In these situations local jurisdictions, while competing for mobile capital, at the same time tax that capital to protect the environment. In this literature, the terms 'environmental standard' and 'environmental regulation' are used interchangeably. See Wilson (1996) and Levinson (2003) for surveys. Our model combines features of models from both areas: on the one hand competition between governments introduces efficiency enhancing incentives; on the other hand the broader environment in which these incentives operate is one of policy-failures that preclude the attainment of a fully efficient equilibrium.

We refer to the difference between a firm's ideal standard level and the level actually set in a country as the 'standard mismatch' for that firm. A key parameter in the model is the 'marginal cost of standard mismatch' ($mcsm$) which parameterizes how a given standard mismatch affects a firm's costs of production. Each firm (being small and behaving non-strategically) chooses its location to maximize profits, taking as given the tax levels and its standard mismatches in the two countries.

Our simple framework yields a surprisingly rich set of equilibrium predictions that depends on the $mcsm$. There are three possible sorts of outcome. (1) If the $mcsm$ is low then fiscal competition leads to an efficient equilibrium outcome (as in Brennan and Buchanan's 1980 model of tax competition). (2) If the $mcsm$ is in an intermediate range then the developed country sets its ES inefficiently high and the developing country becomes a pollution haven; a place where firms that prefer a low ES locate in order to escape the high ES set in the developed country. (3) If the cost of standard mismatch is high then both governments set their ES inefficiently high and, because countries are differentiated by their ES levels, the intensity of tax competition is reduced as well. It is especially interesting that inefficiently high standards can arise in equilibrium, either in the developed country alone (as in 2) or in both countries (as in 3) purely through strategic interaction between governments in their competition for firms and not as a result of attempts by governments to regulate the environment on behalf of consumers/citizens.[8] The precise set of interactions that gives rise to these equilibrium outcomes will be described in due course.[9]

As mentioned above, competition between jurisdictions over standards and taxes has already received some attention in the literature. For example, Oates and Schwab (1998) consider a large number of small jurisdictions who compete in taxes and ESs to attract capital from the world capital market. Markusen, Morey and Olewiler (1995) consider a situation where two jurisdictions compete to attract the plants of a firm.[10] The concern in

---

[8]The issue of governments setting environmental standards too high, in order to dissuade a noxious production facility from locating in their jurisdictions, is referred to as 'not in my back yard' or NIMBY, and brings about a 'race to the top'. In our framework governments do not raise taxes to repel firms since we do not model consumers nor other parties who may be harmed by the hosting of firms. (NIMBY has been studied by Levinson (1999a,b) among others.) Yet we can get a similar outcome to a 'race to the top' through a quite different set of interactions.

[9]Since apart from the sequencing of their policy decisions countries are ex ante symmetrical, firm location decisions are determined solely by the interaction of policy choices with firms' preferences over standards. In equilibrium outcomes (2) and (3), a relatively large share of firms locates in the developing country. We take this to reflect the net flow of firms and capital towards the V4 (see World Bank 1996 page 136) and countries in the emerging markets more generally.

[10]Markusen et al (1995) also consider the possibility of NIMBY.

both settings is with conditions under which competition between governments will lead to a departure from an efficient outcome. These papers make important contributions. Yet as far as we are aware, the situation that we examine here in which some firms inflict pollution externalities on others has not previously been studied in the context of fiscal competition, nor has the focus of attention been the issue of developing country second-mover advantage. We will continue the discussion of how the present paper relates to the literature in Section 5 below.

The remainder of this paper is organized as follows. Section 2 introduces the model. Section 3 solves for the efficient allocation. Section 4 defines strategies and the subgame perfect equilibrium and then characterizes equilibrium in terms of the three cases outlined above. Section 5 places the paper's contribution to the literature and draws conclusions.

2 The Model

The governments of two countries, a developed country, L (for 'leader'), and a developing country, F (for 'follower'), compete over ES levels and taxes in their attempts to induce firms to locate in their respective countries. The governments are assumed to be rent maximizers. There is a set of firms, each of which is able to sell a single unit of a good. The production costs of a firm depend on the level of taxation and the level of the ES in the country where it locates. We will first specify the behavior of firms, and then we will turn to governments. This is the natural sequence of exposition given that we solve for equilibrium using backwards induction.

2.1 Firms

The world price of the unit that each firm sells is p, and each firm pays a private per-unit production cost, c.[11] The tax levied on the firm is τ_L if it locates in L and τ_F if it locates in F. Let the variables l_L, $l_F \in [0, 1]$ denote the ES levels in L and F respectively. The value $s \in [0, 1]$ uniquely identifies a firm and its ideal ES level.[12] The (environmental) *standard mismatch* for a firm s is given by the difference between s and the ES level actually set in the country where the firm locates. The impact of standard mismatch on production costs

---

[11] To increase realism, the price that each firm receives for the good that it sells could be made to vary across firms without affecting the results.

[12] We choose the interval $[0, 1]$ to simplify the exposition. The same qualitative results may be obtained using an arbitrary interval $[a, b]$.

is parameterized by k; we refer to k as the *marginal cost of standard mismatch* (*mcsm*). We can then express the profit function for firm s as follows:[13]

$$\pi\left(s\right) \equiv \begin{cases} p - c - \tau_L - k\left|s - l_L\right| & \text{if the firm locates in } L; \\ p - c - \tau_F - k\left|s - l_F\right| & \text{if the firm locates in } F. \end{cases}$$

To focus the analysis on location decisions, it will be assumed throughout that p is sufficiently high to ensure that all firms make positive profits. Also, $p - c$ will serve as an upper bound for the tax that a government can set.[14]

A firm s makes equal profits in both countries if and only if

$$\tau_L + k\left|s - l_L\right| = \tau_F + k\left|s - l_F\right|,$$

in which case the firm is indifferent between the two countries. If there is a single indifferent firm, $\hat{s}$, then it holds that $\hat{s}$ lies between l_L and l_F. Solving for $\hat{s}$ in this case we obtain:

$$\hat{s} \equiv \hat{s}\left(l_L, \tau_L, l_F, \tau_F\right) = \begin{cases} \frac{\tau_L - \tau_F}{2k} + \frac{l_L + l_F}{2} & \text{if } l_F < l_L \\ \frac{\tau_F - \tau_L}{2k} + \frac{l_L + l_F}{2} & \text{if } l_F > l_L. \end{cases}$$

Firm s may prefer one country, say F, in terms of the tax that it sets; $\tau_F < \tau_L$. But if L's ES is sufficiently close to s (i.e. $\left|s - l_L\right| < \left|s - l_F\right|$) then L can attract s to its country.[15] If there is more than one indifferent firm, then it must hold that for any such firm s, either $s \leq \min\{l_L, l_F\}$ or $s \geq \max\{l_L, l_F\}$. If all firms are indifferent, then $l_L = l_F$ and $\tau_L = \tau_F$. If no firms are indifferent then clearly all firms locate in one country or the other. These cases are treated in the rent functions of the governments defined in Section 2.2.[16]

Three more assumptions are needed to obtain clear-cut solutions for firm locations:

A1. Given taxes and ESs, a firm that is indifferent between the two countries locates in the country where its standard mismatch is lower.

Loosely speaking, A1 captures the idea that firms care more about the persistence of an

---

[13] Here, each firm's profit function is single-peaked over ESs. This constitutes the key difference of our model from other models of fiscal competition, wherein a standard or a public good more generally defined would have a monotonic impact on profits. In Section 5, we outline the differences in outcome from the previous literature generated by our alternative modeling approach. Microfoundations for single-peakedness and some motivating examples are presented in Appendix A1.

[14] We assume p is set sufficiently high relative to c that this bound is never attained.

[15] Firms' location decisions and hence the sizes of the countries, in terms of the measure of firms in each country, are determined strictly by the interaction of policy choices with firms' preferences. Additional features could be introduced to make the model more realistic including, for example, infrastructure and an 'attachment to home' but this would obscure the effects we want to focus on. See Hindriks (1999) for an example of where attachment to home is modeled in the context of tax (versus transfer) competition.

[16] See Appendix A.1 for additional details.

established ES level than the constancy of a given tax level. For $k \leq 1$, any other tie-breaking rule would yield the same results that we obtain. For $k > 1$, any other tie-breaking rule would lead to nonexistence of equilibrium.

A2. If all firms are indifferent between the two countries, then half locate in one country and half locate in the other.

Only if both countries set the same standard and the same tax are all firms indifferent. Assumption A1 is not helpful in this case. Assuming that firms split evenly between the two countries is the most natural tie-breaking rule for this case.[17]

A3. If a government has multiple best responses, it chooses the best response that maximizes its share of firms. If in addition $(l_L, \tau_L) = (0, 0)$, then F sets $\tau_F = 0$.

Again, some tie-breaking rule is needed. We need A3 in two cases. First, for small values of k it is essential for existence of equilibrium. Second, for $k = 1$, firms care equally about low taxes and a close standard match - taxes and standard mismatch are perfect substitutes for them. As a consequence, governments have a continuum of optimal strategies to choose from. Our assumption for this case ensures that the equilibrium outcome for $k = 1$ is the limit of the equilibrium outcome for $k < 1$ as k approaches 1.

The location decisions of firms described above are illustrated in Figure 1.
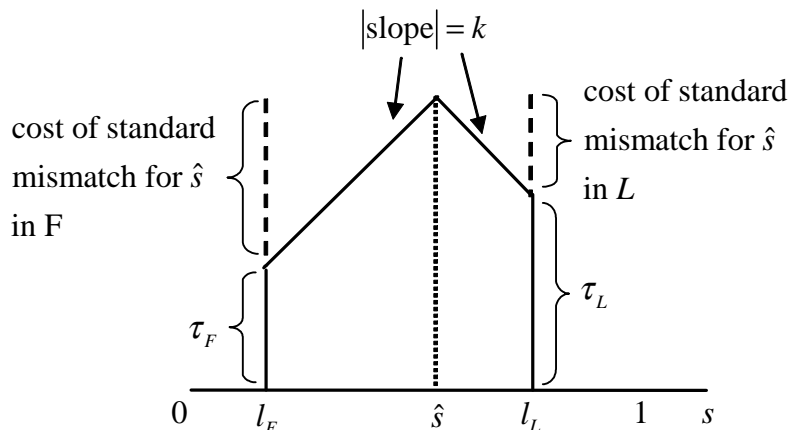
Figure 1

---

[17] A1 and A2 introduce a discontinuity in the share of firms locating in a country. To see this, suppose that $l_F < l_L$ and that $\hat{s} = l_F$. In this case every firm whose ideal standard level is below $\hat{s}$ is also indifferent between F and L. By A1, firms with $s \in [0, \hat{s}]$ locate in F and all other firms locate in L. However, if L lowered its tax by an arbitrarily small amount it could attract all firms to its country. A similar example can be constructed using A2. Admittedly, these discontinuities are not a desirable feature of our assumptions. However, some tie breaking is needed whenever a positive measure of firms is indifferent and *any* tie-breaking rule that we chose would result in some such discontinuities.

Figure 1 is reminiscent of 'Hotelling's umbrella,' and reflects the Hotelling-like features underlying the structure of our model. The figure illustrates the interaction of levels of ESs and taxes set by governments F and L. For ESs and taxes as shown, the point $\hat{s}$ represents the ideal ES level of the indifferent firm $\hat{s}$. For $\hat{s}$, the absolute cost of standard mismatch is lower in L, but the tax in F is lower than the tax in L.

2.2 Governments

Rents are given by tax revenues minus the cost of ES setting. A government's cost of enforcing an ES level $l \in [0, 1]$ is l per firm that is located in its country. Thus the cost of enforcing a given ES is assumed to be proportional to the level of the ES and the number of firms over which it must be enforced.[18] Governments move sequentially.[19] Government F takes l_L and τ_L as parameters and chooses l_F and τ_F to maximize its rents. Discontinuities arise in F's rent function at points where, given L's strategy, F's strategy is such that $\hat{s} = l_F$ or $\hat{s} = l_L$, and additionally when $l_F = l_L$ and $\tau_F = \tau_L$. Below is the rent function for F. The rent function for L is symmetric:

$$r_F\left(l_F, \tau_F; l_L, \tau_L\right) =$$

$$
\begin{cases}
(\tau_F - l_F)\frac{1}{2} & \text{if } \tau_F = \tau_L \text{ and } l_F = l_L \quad \text{Case 1.} \\
(\tau_F - l_F) & \text{if } \tau_F < \tau_L - k\left|l_L - l_F\right| \quad \text{Case 2.} \\
(\tau_F - l_F)\hat{s} & \text{if } \left|\tau_F - \tau_L\right| \leq k(l_L - l_F) \\
& \qquad \text{and } l_F < l_L \quad \text{Case 3.} \\
(\tau_F - l_F)(1 - \hat{s}) & \text{if } \left|\tau_F - \tau_L\right| \leq k(l_F - l_L) \\
& \qquad \text{and } l_F > l_L \quad \text{Case 4.} \\
0 & \text{if } \tau_F > \tau_L + k\left|l_L - l_F\right|. \quad \text{Case 5.}
\end{cases}
$$

---

[18]We are assuming that a higher ES level is more costly due to higher costs of enforcement. Higher ESs call for more controls because firms' incentives not to conform increase and are likely to trigger more court cases. We would not expect our results to change qualitatively if the costs of ES setting were strictly convex instead of linear. Also, if in our model all ES levels were equally costly, existence of equilibrium would become an issue.

[19]If governments set ESs and taxes simultaneously, or if they first simultaneously set ESs and then simultaneously set taxes (both common modeling choices in the literature on fiscal competition) equilibrium in pure strategies would not exist in our model, even if the cost of standard mismatch were quadratic.
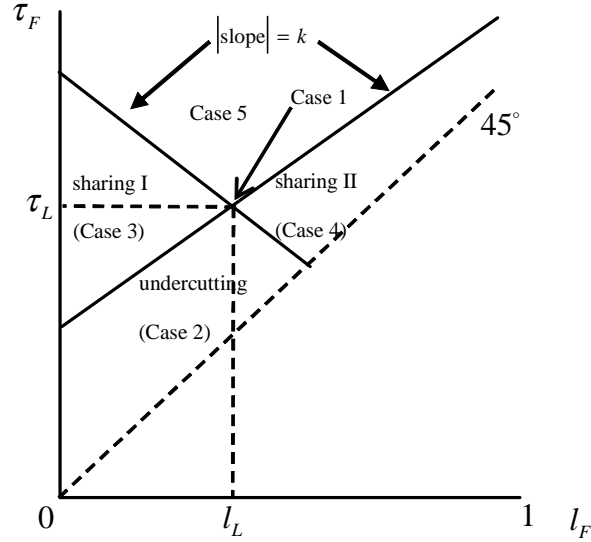
Figure 2

Figure 2 depicts the sets in the strategy space of F corresponding to the different cases of $r_F(\cdot)$. Case 1 arises when both governments choose the same ES and tax levels. By assumption, half of the firms then locate in F. In Case 2, which we will refer to as *undercutting*, the combination of ES levels and taxes induces all firms to locate in F. Cases 3 and 4 arise when strategies result in a positive fraction of firms locating in each of the countries, with F setting a lower ES than L in Case 3 and a higher ES than L in Case 4. We will refer to these third and fourth cases, where firms are shared between the two countries, as *sharing I* and *sharing II*. Finally, Case 5 arises when F chooses its strategy so that it attracts no firms.

3 Efficiency

Within the context of our model, an allocation is efficient if it maximizes the aggregate surplus realized by firms plus the governments' rents. An allocation consists of two ES levels and an assignment of firms to countries, denoted by $(l_F, l_L, \hat{s})$. Formally, the allocation

8

$(l_F, l_L, \hat{s})$ is efficient if it solves[20]

$$\max_{\{l_F, l_L, \hat{s}\}} \int_0^{\hat{s}} (p - c - \tau_F - k\,|l_F - s|)ds + (\tau_F - l_F)\,\hat{s}$$

$$+ \int_{\hat{s}}^1 (p - c - \tau_L - k\,|l_L - s|)ds + (\tau_L - l_L)\,(1 - \hat{s})$$

$$s.t. \quad l_F \in [0, 1]\,, \, l_L \in [l_F, 1]\,, \text{ and } \hat{s} \in [0, 1].$$

The integrals are the profits of firms that are allocated to the two countries. The other two terms are the rents of the two governments. The problem can be simplified to

$$\min_{\{l_F, l_L, \hat{s}\}} \int_0^{\hat{s}} k\,|l_F - s|\,ds + l_F\hat{s} + \int_{\hat{s}}^1 k\,|l_L - s|\,ds + l_L\,(1 - \hat{s})$$

$$s.t. \quad l_F \in [0, 1]\,, \, l_L \in [l_F, 1]\,, \text{ and } \hat{s} \in [0, 1].$$

Thus the efficient allocation minimizes the sum of the aggregate costs of standard mismatch and the costs of ES setting. We use superscript e to denote an efficient allocation. To express dependencies on k, we write $l_F^e(k)$, $l_L^e(k)$, and $\hat{s}^e(k)$.

It is immediate that, if $k < 1$, the set of efficient outcomes is given by $l_F^e(k) = 0$, $l_L^e(k) = 0$, and $\hat{s}^e(k) \in [0, 1]$. That is, for $k < 1$ it is efficient to set a zero ES with the share of firms that locates in each country being indeterminate. Even for the firm $s = 1$, it is more efficient to incur the costs of standard mismatch, k, than to pay for a positive ES level l that would lower mismatch costs: $k < l + k\,(1 - l) = k + l\,(1 - k)$. If $k = 1$, any allocation for which $l_F^e = 0$ and $l_L^e = \hat{s}^e \in [0, 1]$ is efficient. In addition, for $l_F^e = l_L^e = 0$ any $\hat{s}^e \in (0, 1]$ is efficient as well. Since the *mcsm* and the marginal cost of enforcing the standard for an additional firm are equal if $k = 1$, there exists a continuum of efficient allocations.

For $k > 1$, solving the minimization problem above yields the efficient allocation:

$$l_F^e(k) = \frac{k - 1}{4k};$$

$$l_L^e(k) = \frac{3k - 1}{4k};$$

$$\hat{s}^e(k) = \frac{1}{2}.$$

The efficient standard levels are increasing in k. Figure 3 illustrates the efficient ES levels and the allocation of firms to countries depending on k for the case $k > 1$.

---

[20] If it is efficient that the two countries set different standard levels, it does not matter for the efficiency of the allocation whether F or L sets the higher standard. Here, we pose the problem so that L sets a standard not lower than F. Since the roles of F and L can be exchanged, the results in this section are unique only up to a relabeling of countries.
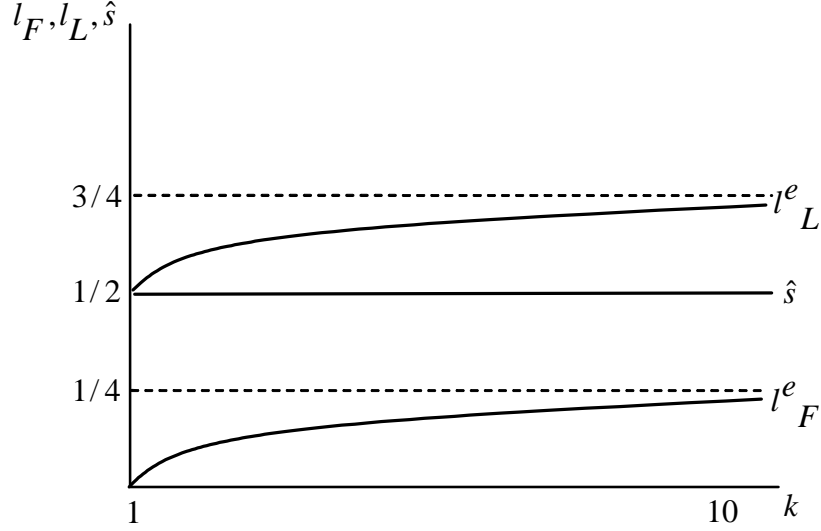
Figure 3

Given the Hotelling features of our underlying model, one might have expected the efficient solution to have the form $l_F^e(k) = \frac{1}{4}$ and $l_L^e(k) = \frac{3}{4}$ familiar from Hotelling (1929). In our model the efficient levels of enforcement are lower, starting at just above $l_F^e(k) = 0$ and just above $l_L^e(k) = \frac{1}{2}$ respectively for $k \to 1$ (from above) and converging towards $l_F^e(k) = \frac{1}{4}$ and $l_L^e(k) = \frac{3}{4}$ respectively as k becomes large. To understand why our efficient ES levels are lower than they would have been in a direct application of Hotelling, recall that in our model one has to take into account the costs of enforcing the ES for each firm assigned to a country as well as the costs of standard mismatch. If our model were a direct application of Hotelling then the level of the ES would not have affected its cost of enforcement. Efficient ES levels in our model approach the efficient levels that would have arisen in a direct application of Hotelling's model as k becomes large because the cost of standard mismatch becomes large relative to the cost of enforcement. Finally, as in Hotelling's model, in our model the share of firms between countries is equal. This efficient solution will serve as a benchmark against which to compare the equilibrium outcome.

4   Competition over Environmental Standards and Taxes

In this section, our approach will be to first define equilibrium and then state our main theorem in which equilibrium is characterized. After that, we will undertake a diagrammatic discussion of the properties of equilibrium. A sketch of the proof is presented in the appendix. For a full formal derivation of equilibrium see the appendix.

As mentioned above, ES provision and tax setting are modeled as a two-stage game. The sequence of events is as follows. Government L sets its ES level and tax and then, observing L's choices, Government F sets its ES level and tax. Taking government policies as given, firms then make location decisions to maximize profits. As usual, a *subgame perfect Nash equilibrium* is a strategy profile with the property that the governments' strategies constitute a Nash equilibrium in every subgame of the game.

A strategy for Government L is a pair consisting of an ES level and a tax. A strategy is *feasible* if the tax is high enough to cover the cost of ES setting.[21] Formally, the set of feasible strategies is

$$S_L = \{(l_L, \tau_L) \in [0, 1] \times [0, p - c] \mid \tau_L \geq l_L\}.$$

A strategy for Government F is a mapping that assigns a pair, consisting of an ES level and a tax, to each possible strategy choice made by Government L in the first stage of the game. Formally, this mapping is described by $f : S_L \rightarrow [0, 1] \times [0, p - c]$ where $f(l_L, \tau_L) = (l_F, \tau_F)$. Let $\mathcal{F}$ be the set that contains all such mappings. The set of feasible strategies for Government F consists of those members of $\mathcal{F}$ with the property that tax revenue covers the cost of the associated ES level; that is,

$$S_F = \{f \in \mathcal{F} \mid \text{for all } (l_L, \tau_L) \in S_L, \, f(l_L, \tau_L) \text{ satisfies } \tau_F \geq l_F\}.$$

We are interested in the pure strategy subgame perfect Nash equilibrium of the game, which can be viewed as a Stackelberg game.[22]

Definition 2. *A pure strategy subgame perfect Nash equilibrium in taxes and ES levels is a pair of strategies $((l_L^*, \tau_L^*), f^*)$ such that*

1. $(l_L^*, \tau_L^*) \in S_L$ *is a best response to* f^*.

2. $f^* \in S_F$ *and* $f^*(l_L, \tau_L)$ *is a best response to* (l_L, τ_L) *for all* $(l_L, \tau_L) \in S_L$.

With the structure of the model in place and equilibrium defined, we are now ready to state our main theorem which characterizes equilibrium.

---

[21]Thus we make the simplifying assumption that there are no other sources of government revenue and no international capital market that governments can tap. We do not think that allowing such a possibility would change our results, since governments make positive rents in equilibrium.

[22]It will be assumed throughout that mixed strategies in tax rates are not available to governments. This is generally deemed to be an acceptable assumption in the applied literature on policy setting in a perfect information environment.

Theorem 1. *The outcome of the subgame perfect equilibrium.*[23]

The subgame perfect equilibrium is as follows.

a. *(Efficient outcome) If $k \leq \frac{1}{3}$, both L and F set the minimum ES level and set zero taxes. Firms split equally between the two countries; that is, $(l_L^*, \tau_L^*) = (0,0)$ and $(l_F^*, \tau_F^*) = (0,0)$, and $\hat{s}^* = \frac{1}{2}$.*

b. *(Pollution haven) If $\frac{1}{3} < k \leq 1$, the differentiation in ES levels between the two countries is high; L sets an ES close to the maximum level and F sets a zero ES level. Both governments set taxes that lead to positive rents, and rents are always higher for F than for L. The majority of firms locates in F. Specifically, it holds that $l_L^* \geq \frac{8}{9}$, $\tau_L^* \in (l_L^*, 2l_L^*)$, and $l_F^* = 0$, $\tau_F^* \in \left(\frac{2}{3}, \frac{4}{3}\right)$, and $\hat{s}^* > \frac{2}{3}$.*

c. *(Excessive ESs) If $k > 1$, the ES level is above $\frac{1}{2}$ in both countries, with L setting a higher ES than F. The ES levels do not vary with k. Both governments make positive rents, requiring firms to pay more than twice the cost of ES provision. F sets a higher tax than L and earns higher rents. Two-thirds of the firms locate in F, and every firm with a strictly higher ideal ES level than set in F locates in L. Specifically, it holds that $l_L^* = \frac{8}{9}$, $\tau_L^* = \frac{4}{3} + \frac{4}{9}k > 2l_L^*$, and $l_F^* = \frac{2}{3}$, $\tau_F^* = \frac{4}{3} + \frac{2}{3}k > 3l_F^*$, and $\hat{s}^* = \frac{2}{3}$.*

Figure 4 shows the equilibrium ES and tax levels set in the two countries depending on k. The solid curves represent the ES and tax set by L and the dashed curves represent the ES and tax set by F. As the figure shows, the subgame perfect ES and tax levels differ considerably across the three regions of k: In case (a) a small k leads to an efficient outcome with both countries setting the minimum ES level and taxes set at the same level in both countries at $\tau_L^* = \tau_F^* = 0$; in case (b), for k in an intermediate range, there is almost maximum differentiation in ESs, with F setting the efficient ES while L sets its ES too high and sets a higher tax than F; in case (c), for large k, there is some differentiation but it is substantially smaller than for k in the intermediate range, and both countries set inefficiently high ESs while F now sets a higher tax than L.

---

[23]The theorem is restated in Appendix A.3 with formulae for all the equilibrium values shown explicitly.
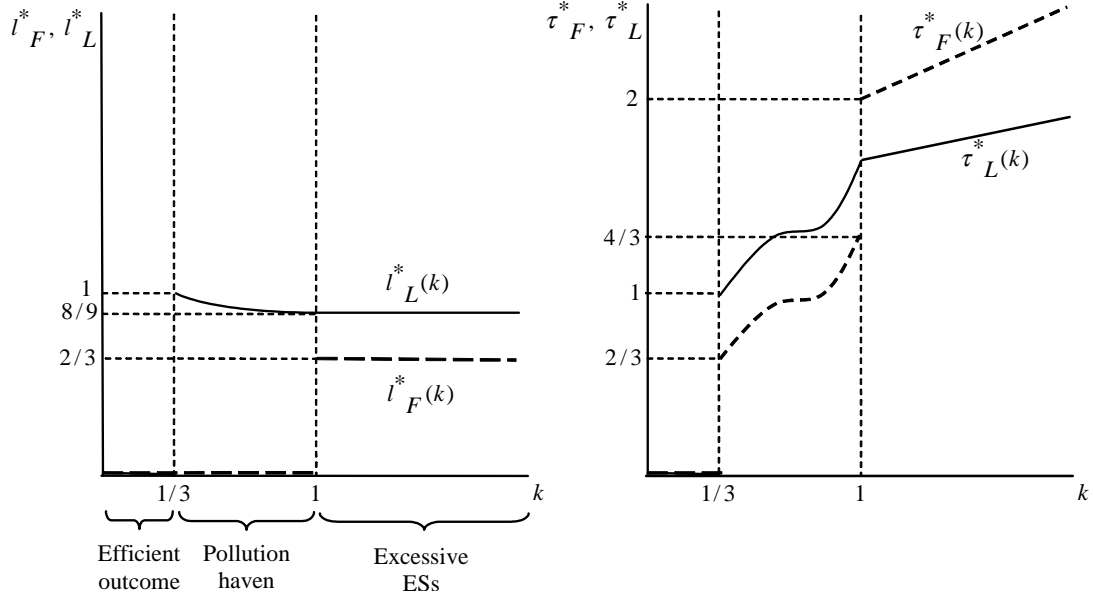
Figure 4

The common characteristic of equilibrium across all levels of k is that F attracts at least as many firms as L. Also, for $k > \frac{1}{3}$, governments are able to extract rents. This arises as a result of the monopolistic power that each government has over location within its country. Each firm must locate in one country or the other in order to produce, and the government of the country where it does locate is able to exploit its resultant power when setting taxes. In addition, F, who sets a lower ES, makes more rents because it both attracts more firms and makes more rents per firm.

In the following we discuss in more depth the intuition behind cases (b) and (c), providing specific details about the features of our model that drive them. The reason that we do not discuss case (a) further is because its logic is familiar from Brennan and Buchanan (1980). Each firm regards it as relatively unimportant to obtain a close standard match; thus all competition occurs in taxes, which brings about an efficient outcome.

Turning to case (b), with k in an intermediate range, it is a dominant strategy for F to set its ES at zero. Government L can extract some rents (because k is not 'too small'), but only by differentiating itself substantially from F (because k is not 'too large'). But L can only differentiate itself by setting its standard at a sufficiently high level, with close to maximum differentiation between the two countries.
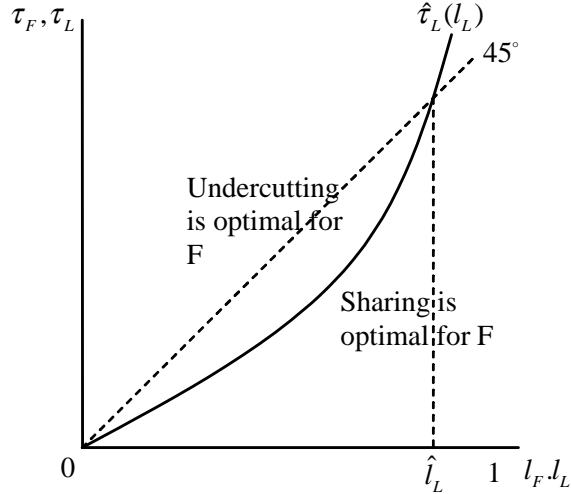
Figure 5

Figure 5 helps to understand the strategic interaction between the two governments. It depicts what we call the *sharing tax limit*, denoted by $\hat{\tau}_L(l_L)$. For standard level l_L, the sharing tax limit $\hat{\tau}_L(l_L)$ is the corresponding maximum tax that L can set without inducing F to choose an undercutting strategy, as a result of which L would loose its entire tax base. Figure 5 shows what the sharing tax limit would typically look like if $k \in (1/3, 1]$. To make positive rents, L must also set a tax strictly above the cost of standard setting (i.e. above the dashed line). To achieve both objectives simultaneously, L must choose a standard level at least as large as $\hat{l}_L$. For $l_L \geq \hat{l}_L$, the relatively small distance between any value of l_L and the corresponding value of $\hat{\tau}_L(l_L)$ allows L to make only low rents compared to F which is free to maximize the trade-off in tax setting between its share of firms and the tax that it levies on each. This provides the essential intuition behind why F obtains a second-mover advantage.

In case (c), as in case (b), a sharing tax limit imposes an upper bound on the tax that L can set while still making positive rents. However, the difference in case (c) is that both countries set inefficiently high ES levels.[24] Because of the higher value of k, a higher standard mismatch hurts firms' profits more than a higher tax and so it is no longer a dominant strategy for F to set a zero ES level. Instead, Government F chooses an ES level close to L's ES level (but not arbitrarily close as this would negatively affect F's ability to extract rents). Government L chooses a high ES level to induce sufficient differentiation and to leave F large sharing rents which in turn allows L to set a relatively high tax without

---

[24]Recall from Section 3 that the efficient outcome calls for the countries offering up to, respectively, 25% and 75% of the maximum standard level.

14

inducing F to undercut. This incentive to 'compete up' ES's leads both governments to set them too high in equilibrium.
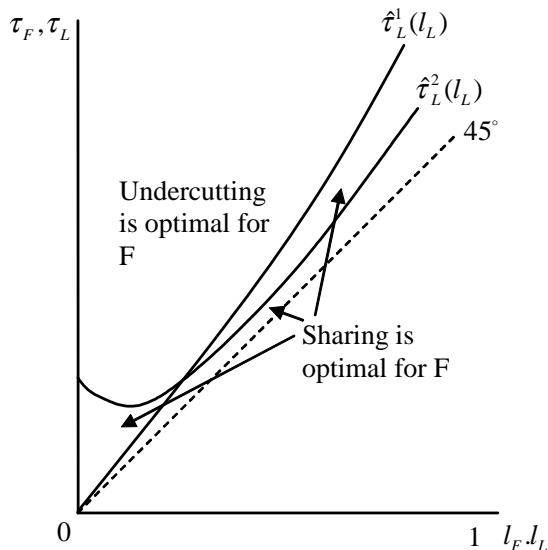

Figure 6

The sharing tax limit for case (c) is depicted in Figure 6. The sharing tax limit is the upper envelope of the two curves in the figure, that is $\hat{\tau}_L(l_L) = \max\left\{\hat{\tau}_L^1(l_L), \hat{\tau}_L^2(l_L)\right\}$, where $\hat{\tau}_L^1(l_L)$ is the tax up to which F chooses a sharing I strategy rather than an undercutting strategy and $\hat{\tau}_L^2(l_L)$ is the tax up to which F chooses a sharing II strategy rather than an undercutting strategy. If L sets a sufficiently low ES level and a sharing tax, then F has an incentive to share with a higher ES level because this yields higher rents than sharing with a lower ES level; this holds in reverse if F sets a sufficiently high ES level. Inducing F to share with a lower ES level has the advantage that the decreasing cost of standard setting gives F an incentive to choose its standard not too close to l_L. Moreover, as suggested by Figure 6, if L sets a relatively high ES level this allows it greater scope to set a relatively high tax without inducing undercutting. These combined effects enable L to make higher rents than if it set l_L relatively low and induced F to set l_F relatively high. But its rents are still lower than those of F, so case (c) exhibits developing country second-mover advantage as well.

5   Relation to the Literature and Conclusions

We began this paper by noting concerns in policy circles that developing countries resembling those of recent entrants to the EU may, under certain circumstances, have a second mover

15

advantage in setting ESs and taxes. We then set out a formal framework which makes precise a set of circumstances under which such a second-mover advantage may arise. Three possible predictions were made about the outcome of fiscal competition when the public good in question is an ES. The particular prediction that emerges in equilibrium depends on the *mcsm*. The model focuses on the interplay between governments' incentives to manipulate policy - ESs and taxes - in order to maximize rents and firms' incentives to locate where these policies have the most favorable impact on their profits. The key point is that the government of the developed country wants to avoid inducing the developing country to undercut because that implies losing the entire tax base and hence all rents. If the *mcsm* is low, then ESs are not important enough to firms for governments to be able to use them strategically. In this case, the forces of tax competition envisaged by Brennan and Buchanan dominate, and the outcome is efficient. If the marginal cost of standard mismatch is high enough, the developed country government successfully induces sharing by setting a sufficiently high ES relative to the tax. A proportion of firms then find it beneficial to locate in each country. Governments are able to use policy to make rents, and the resulting outcome is inefficient in that either the developed country government or both governments set ESs too high.

We will now place the paper's contribution to the literature, starting with the literature that follows Tiebout (1956) and then moving on to consider the literature that parallels fiscal federalism and standard tax competition. As in the literature that follows Tiebout, governments in our model are rent (or profit) maximizing but are constrained by competition. For example, Fischel (1975) and White (1975) share with the present paper the assumption that there is variation over firms' preferences for standards. In keeping with our model, there are no cross-border externalities. In contrast to our model, Fischel (1975) and White (1975) both assume that individual firms can be targeted for transfers and there is 'free entry' of jurisdictions, none of which has sufficient market power to extract rents from firms.[25] As a result, within such a setting, an efficient outcome can be demonstrated in which firms 'vote with their feet.' In our model firms cannot be targeted for transfers. There is policy failure in the sense that once the policies are set they cannot be altered. And there are only two jurisdictions. It is interesting to note that none of these differences in modeling approach matter for the achievement of efficiency providing that the *mcsm* is sufficiently small. It is

---

[25] Our focus in this paper on the issue of second-mover advantage by developing countries dictates that the policy-environment is pitched at the level of national governments whereas in much of the prior literature the focus is on jurisdictions within a country or federation. In addition, the range of policy options that we consider are more limited than under federalism, mirroring more closely an international setting.

only when the *mcsm* becomes sufficiently large for governments to compete for firms using ESs that a divergence from efficiency arises. Given the focus on efficient outcomes in this first area of the literature and the fact that the focus tends to be on cooperative frameworks, there is limited scope in such a setting for exploring the type of second-mover advantage that is our focus in the present paper.

Let us turn now to the second area of the literature, that parallels the literature on fiscal federalism and tax competition.[26] A common feature across the two areas is that jurisdictions benefit when total tax revenue is larger, yet at the same time each jurisdiction attracts tax base by lowering taxes and/or environmental standards. In this second area of the literature, as in our model, policy makers cannot target firms for direct transfers. In contrast to the first area and also in contrast to our model, owners of the mobile resource (here capital) do not care about ESs and, seeking the highest return, tend to move their capital away from a jurisdiction if required to foot the bill for an ES. Finally, in this second area of the literature governments are benevolent and use policy to maximize the welfare of their citizens, usually including consumers, which contrasts with the first area of the literature and with our model.

In spite of the differences between the models in the second area and our model, the forces of competition between governments can operate in a similar way. This is seen most clearly by comparing the model of Markusen et al (1995) to ours. Recall that in Markusen et al (1995), two jurisdictions compete for the plants of a firm using two pollution taxes, one on domestic production and one on exports. The benchmark situation is where the firm locates all its production facilities at home. Yet providing that transport costs are high enough and plant set-up costs are low enough, the foreign government can undercut the home government, much as in our model, to get some or all production to locate in its jurisdiction.

In contrast to our model, however, the idea of 'second-mover advantage' cannot be motivated in models from this second area of the literature. The outcome tends towards a situation where both jurisdictions obtain the same level of welfare. Levinson (1997) highlights this outcome by rewriting the model of Markusen et al so that the monopoly rents are earned locally to where a plant locates. This set-up provides the clearest setting in which to see that the country that hosts the firm at the outset has an incentive to act in the manner of a limit-pricing monopolist, 'limit-taxing' the other country, by setting taxes just low enough

---

[26]This literature builds on an earlier literature, initiated by Cumberland (1979, 1981), that is concerned with competition between governments over environmental standards alone.

that the firm is indifferent between locating its plants in one country or two and welfare is the same in both countries.

The effect of competition between jurisdictions is similar in Davies and Ellis (2007) although the set-up is somewhat different. In their model, each firm benefits from the provision of a public good by other firms. This public good could be a reduction of polluting emissions from the production process (although Davies and Ellis' discussion is not limited to environmental externalities). The standard in their model mandates all firms who locate in that jurisdiction to provide the public good at the efficient level while a tax break transfers rents to firms and induces them to locate there. Davies and Ellis straddle the two areas of the literature on taxes and standards in that firms benefit from standard provision but a policy imperfection induces competition between the two jurisdictions. These combined features are crucial in driving the result that the outcome is efficient in their framework, with firms providing the standard at the efficient level and all rents being transferred to them via the competition in taxes.[27] Their symmetric model gives rise to a symmetric equilibrium, with neither government obtaining an advantage over the other.

It may be helpful to see the effects that motivate our results by direct comparison to the broader literature on fiscal competition. We are comparing the role of competition over taxes and public goods more conventionally defined against the role of competition over taxes and ESs in an environment where firms exert externalities on one another. The novel feature of competition in our model is that the developed country is compelled to set its standard sufficiently high that it does not induce the developing country to undercut, and that this effect is strong enough to give the developing country a second-mover advantage. This leads at least the developed country and sometimes the developing country as well to set the standard too high. So our aim is to establish the novelty of this effect through comparison to the rest of the literature that has demonstrated over-provision of the public good.

In a standard model of fiscal competition, Keen and Marchand (1997) consider a setting where the composition of public good provision matters. A jurisdiction can attract capital by increasing its productivity through a shift in spending from 'public consumption goods' such as parks and art galleries to 'public investment goods' such as road and communications networks. In this setting, competition in public good provision, which raises the return to capital that locates there, works in much the same way as tax competition.

---

[27]Davies and Ellis (2007) discuss further how their efficiency result can be sensitive to the set-up of their model.

Thus, if the public good-capital complementarity is sufficiently strong, the equilibrium can exhibit over-provision of the public good in equilibrium. Although the setting of Keen and Marchand's discussion invites a consideration of agglomeration externalities they leave these forces aside. Focusing on the forces of competition over public goods alone, there seems no reason to suppose that one jurisdiction or another should derive an advantage from this kind of competition, and ex ante symmetrical jurisdictions give rise to ex ante symmetrical outcomes for the respective jurisdictions in this setting.[28]

Baldwin and Krugman (2004) do focus on agglomeration externalities. In their model, scale economies induce economic activity to become concentrated. Competition takes place over taxes and in equilibrium all of the (mobile) capital locates in only one jurisdiction. As in our framework, one government leads in policy-setting and the other follows. Like in Markusen et al, one jurisdiction has all the capital to begin with, and as a result it can limit-tax the other jurisdiction and keep all the capital for itself. At the same time, since agglomeration creates rents for firms that can be taxed, the forces of tax competition can lead to excessive levels of taxation (while in our model taxes are welfare neutral as they simply transfer rents from firms to the government). Unlike in Markusen et al, the government of the jurisdiction where all capital locates to begin with can skim off some of the agglomeration rents in the form of tax revenue, giving it a *first*-mover advantage.[29]

In our model, the presence of the ES and firms' single-peaked profits over the ES allow governments to differentiate themselves from one another. As in Baldwin and Krugman, the second-moving country shares firms only if it cannot do better by undercutting. However, for

---

[28]Brueckner (2000) considers Tiebout/tax competition in an environment where firms' public good requirements vary, and shows that firms whose requirements are similar sort themselves efficiently across jurisdictions. The model of the present paper shares the feature of Tiebout-tax competition that there is variation in firms' public good requirements. Another common feature is that governments' objectives are entirely self-serving in that they are profit/rent maximizing but are constrained by competition. In contrast to Tiebout/tax competition where there is no policy failure, the policy-failure in our model does allow governments to have market power and this underpins the difference in outcome that efficiency is not achieved in equilibrium.

[29]Other papers where one jurisdiction is able to limit-tax the other include Black and Hoyt (1989) and Haufler and Wooton (1999). Kind, Knarvik and Schjelderup (2000) and Boadway, Cuff and Marceau (2004) study tax competition in the presence of scale economies. Public investment goods do not usually play a role in such models. An exception is Justman, Thisse and van Ypersele (2002), who investigate the idea that under fiscal competition regions can segment the market for industrial location by offering infrastructure services that are differentiated by quality. They identify a fiscal agglomeration property, which motivates an asymmetric equilibrium. In a related setting, Zissimos and Wooders (2007) show that if governments are able to differentiate the public good that they provide then this reduces the intensity of (beneficent) tax competition and thus reduces efficiency. In these papers, if any government has an advantage in the sequence of play it is the first-mover.

two reasons undercutting happens in a more aggressive fashion in our model. First, there is no immobile tax base in the second-moving country. Second, in our framework a marginally reduced tax suffices to undercut (while in Baldwin and Krugman, to attract the mobile tax base, the second-moving country has to compensate it for its loss in agglomeration rents as the entire mobile tax base is initially located in the first-moving country). To induce sharing in the presence of the aggressive undercutting behavior featured in our model, the first-moving country has to leave most of the rents to the second-moving country.

Our model shares similarities with Bucovetsky (2005) as well. He examines an interesting situation that blends key features of previous models; many jurisdictions may compete for mobile capital in a setting where agglomeration externalities are important. To focus on the special features of competition over public investment goods, or 'public inputs' as he refers to them, he abstracts from tax competition entirely. If agglomeration externalities are sufficiently strong, the efficient solution has all of the mobile factor locating in one jurisdiction. Yet competition among jurisdictions can lead to a Nash equilibrium in which the mobile factor locates across more than one jurisdiction, in contrast to Baldwin and Krugman and similar to our model. Competition between jurisdictions can lead to over-provision of the public good as in our model. A nice feature of Bucovetsky's model which ours does not share is that comparative statics can be carried out whereby the economies of scale can be increased by an increase in the degree of substitutability between the goods that jurisdictions produce, potentially bringing about a decrease in the number of jurisdictions that provide public investment goods in equilibrium. Our model is different in the specific mechanism by which ESs are set excessively high. In Bucovetsky over-provision of the public input arises as jurisdictions fail to take into account the negative externalities that an increase in the public input inflicts on other jurisdictions. In our model there are no such cross-border externalities. Instead, the developing country's second-mover advantage leads the developed country to set an inefficiently high ES to induce less aggressive competition by the developing country.

It is worth drawing parallels between our work and the large literature, primarily in the field of international trade, that has focused on pollution havens. The pollution haven hypothesis is that, as economies open up to each other, dirty industry will tend to become concentrated in the country with the weakest ESs. Standard international trade theory provides a natural explanation for this, which explains why it forms the cornerstone of the

main explanation that is put forward for the possible existence of pollution havens. The idea is that, all else equal, thinking of pollution as an 'input' to the production process, lax ESs are a source of comparative advantage since they make the opportunity cost of pollution relatively low. Antweiler, Copeland and Taylor (2001) construct a model around this idea and present cross-country empirical evidence that provides some support for the existence of pollution havens (also see Taylor 2004). More recent empirical work calls into question the existence of pollution havens on the basis that the pollution content of trade flows do not appear to support the predictions of the trade model; see Ederington, Levinson and Minier (2004). Part (b) of our Theorem 1 is helpful in this regard since it presents an alternative strategic motivation for the existence of pollution havens in developing countries based on the feature of our model that the developing countries we focus on are able to respond quickly to the policies of developed countries and hence undercut them if it is profitable to do so.

Inevitably, the theoretical framework developed here simplifies the situation in a number of key respects. For example, to keep the analysis manageable we have not explicitly treated consumers in our analysis and we have restricted the number of countries to just two. A promising direction for future research would be to extend our model to give consumers a more prominent role. One potential limitation to our conclusions is that the government in the developing country does not set ESs 'too low.' While it seems reasonable to argue that developed countries may set ESs too high, a concern is that developing countries actually set their ESs too low from the perspective of consumers. The introduction of consumers to the model could make it possible for ESs to be set too low in the developing country.

Another promising direction for future research would be to ask how robust our results would be to the introduction of a larger number of countries to the model. From our analysis of the present framework it is not obvious how the outcome would be changed by the introduction of more countries. One conjecture would be that the nth country to move would always have the greatest advantage, with prior countries being constrained by those that would set policy subsequently. A different conjecture about the outcome would be that only two countries could make positive rents and that the presence of more countries would be irrelevant. If the analysis of a larger number of countries turned out to be analytically intractable then it might be possible to obtain characterizations through numerical simulation.

Finally, a question that could be addressed in the future is whether incentives exist for

governments to coordinate/harmonize policy within our framework. Under perfect collusion in our model, governments would simply agree that neither of them would set a positive ES level and they would set taxes at the level of profits, thereby extracting all surplus. Such an outcome would be efficient in our framework in the case where $k \leq 1$ because in that case the efficient outcome has zero ESs; for $k > 1$ the efficient outcome does have a positive level of ES provision. However, such perfect collusion would require a strong enforcement mechanism and, in the absence of an international enforcement body, the incentives to break such an agreement might be overwhelming. This may explain why in practice proposals for collusion have tended to be weaker, entailing for example the introduction of minimum ESs. A surprising implication of our framework is that it is not in the interest of the developed country to introduce a binding minimum ES. The reason is that the developed country benefits from being able to differentiate itself from the developing country and putting in place a minimum ES would limit the scope for doing so. Thus our model presents a possible way of understanding situations in which minimum ES levels have been called for but none have actually emerged. The question of policy coordination/harmonization appears to raise issues that warrant further investigation in future research.

A   Appendix

A.1   Single-peaked Profits

The following describes a model that generates profits that are single-peaked in the standard. Suppose that every standard level $l \in [0, 1]$ results in two types of costs for firms: a cost of conforming to the standard and an externality cost arising from pollution by other firms when the standard is set at that level. More specifically, assume that

$$c(l, s) = (1 - s)l^2 + s(1 - l)^2,$$

where $s \in [0, 1]$ is uniformly distributed. A firm of 'type' s incurs a cost of $(1 - s)l^2$ from conforming to the standard l and incurs a negative externality of $s(1-l)^2$ when the standard level enforced is l. Note that this negative externality decreases in the standard.

For motivation, suppose that the standard in question is water quality. The government achieves a certain quality level by setting maximum pollution levels of waste water. If a firm produces polluted waste water as part of its production process, higher levels of water quality are associated with higher costs for that firm because the waste water has to be

cleaned before being sent back to the water system. Firms with low levels of s are those firms for which the cleaning process is particularly expensive, either because their waste water is highly polluted or because they produce large amounts of polluted water. If water of a certain quality is an input to a firm's production process (e.g. in a food processing industry or in the water industry), higher levels of water quality are associated with lower levels of production costs. If other firms pollute the water system, the firm has to purify the water itself before using it. Again, the parameter s captures different requirements for clean water across firms. The same formal reasoning applies to other environmental standards such as standards associated with noise regulation, air or soil quality.

For the cost function $c(l, s)$, a type s firm's cost is minimized at the standard level $l^* = s$, i.e. its profits are single-peaked in the standard with the peak at $l = s$. Moreover, since s is uniformly distributed over $[0, 1]$, so are the peaks, just as we assume in the paper. If the firm locates in a jurisdiction where the standard level deviates from l^* by x, the firm's costs go up by x^2 compared to its costs at l^*. Notice that the cost of standard mismatch is convex while it is linear in the model we employ in the paper. We assume a linear cost for analytical tractability and believe that the convexity of the cost would not change the qualitative nature of our results.

Note that the negative externality depends on a firm's type and is increasing in the standard level, but does not vary with the share of firms located in the country. Thus, we assume that the cost of a filter to clean water depends on the firm's specific requirements for clean water (captured by s) and on the standard set in the country. However, once the filter is installed it can purify water at no marginal cost and whether there are several or only one firm upstream who pollute the water does not affect the cost of cleaning the water.[30] If instead the negative externality was $yb(1 - l)^2$, with y being the share of firms located in the country in question, firms' profits would still be single-peaked with peak $l = \frac{yb}{1 - b(1-y)}$. The distribution of the peaks, however, would depend on the share of firms located in a country, which is not consistent with the model in the paper and considerably complicates the analysis.[31]

---

[30] We need to make the additional innocuous assumption that in each country there is an arbitrarily small fraction of immobile firms polluting the environment so that there always exist some firms in each country that generate negative externalities.

[31] It is worth emphasizing that the negative externality could also be one that a firm incurs more indirectly because it has to compensate its workers for negative externalities they bear. For example, a firm that locates in a jurisdiction with poor air and water quality might have to compensate its workers to work in such an environment.

A.2 Indifference Set

The following is an application of the approach taken by d'Aspremont, Gabszewicz, and Thisse (1979) to the present setting. Given $(l_L, \tau_L, l_F, \tau_F)$, there may be more than one firm that is just indifferent between the two countries. To deal with this possibility, we define the *indifferent set* of firms and denote it by $I(l_L, \tau_L, l_F, \tau_F)$. If the Indifferent Set is not a singleton, a tie breaking rule is needed to determine where indifferent firms locate. With two exceptions, the indifferent set $I(l_L, \tau_L, l_F, \tau_F)$ will be a singleton set, i.e., $\hat{s}(l_L, \tau_L, l_F, \tau_F)$ is the only member of $I(l_L, \tau_L, l_F, \tau_F)$. The two exceptions are as follows.

(1) Suppose that $l_F < l_L$ so that F sets a lower standard than L. For s satisfying $s = l_F$, if $s \in I(l_L, \tau_L, l_F, \tau_F)$ then for all $s' < s$, it holds that $s' \in I(l_L, \tau_L, l_F, \tau_F)$. To see this, first note that for firm $s \in I(l_L, \tau_L, l_F, \tau_F)$, $s = l_F$, the extent to which the tax in F exceeds the tax in L exactly matches the cost of standard mismatch in L, i.e. $\tau_F - \tau_L = k(l_L - l_F)$. Compared to the costs the firm $s = l_F$ has in F and L, respectively, a firm $s < l_F$ has an additional cost of standard mismatch of $k(l_F - s)$ in either F or L, implying that those firms must be indifferent as well and that $I(l_L, \tau_L, l_F, \tau_F) = [0, l_F]$. By analogous reasoning, if firm $s = l_L$ is indifferent, then $I(l_L, \tau_L, l_F, \tau_F) = [l_L, 1]$. The case $l_L < l_F$ is symmetric.

(2) Suppose that $l_F = l_L$; in this case a firm's choice of location is determined by taxes. If $\tau_F = \tau_L$ then all firms are indifferent and again $I(l_L, \tau_L, l_F, \tau_F)$ is not a singleton set but equals $[0, 1]$.

Note that it might also be the case that no firm is indifferent. For example if $l_L = l_F$ and $\tau_F \neq \tau_L$, all firms prefer whichever country sets the lower tax; consequently $I(l_L, \tau_L, l_F, \tau_F)$ is the empty set. More generally, whenever one country undercuts the tax of the other country by more than the cost of the standard difference between the two countries, the indifferent set will be empty.

A.3 Derivation of Theorem 1 and Proofs

A.3.1 The developing country's best response function

This subsection provides a characterization of F's best response, and this is followed by a characterization of L's best response in the subsection that follows. To find Government F's best response to a given strategy of L, we proceed in two steps. First, we maximize F's rents

separately over the three response subsets, *sharing I*, *sharing II*, and *undercutting*.[32]

Government F's optimization problem.

(a) Maximize rents over sharing I

$$\max_{\tau_F, l_F}(\tau_F - l_F)\hat{s}$$

s.t.

$$l_F \in [0, l_L)$$

$$\tau_F \in [l_F, p - c]$$

$$\tau_F \in [\tau_L - k(l_L - l_F), \tau_L + k(l_L - l_F)]$$

(b) Maximize rents over sharing II

$$\max_{\tau_F, l_F}(\tau_F - l_F)(1 - \hat{s})$$

s.t.

$$l_F \in (l_L, 1]$$

$$\tau_F \in [l_F, p - c]$$

$$\tau_F \in [\tau_L - k(l_F - l_L), \tau_L + k(l_F - l_L)]$$

(c) Maximize rents over undercutting

$$\max_{\tau_F, l_F}(\tau_F - l_F)$$

s.t.

$$l_F \in [0, 1]$$

$$\tau_F < \tau_L - k|l_L - l_F|.$$

Second, given the solutions to (a), (b), and (c), the best response is found by comparing the maximized rents across the three sets of possible solutions. To keep track of the different kinds of sets characterizing our results, we introduce the following notation. The responses that maximize $r_F(l_F, \tau_F; l_L, \tau_L)$ over undercutting, sharing I, and sharing II are denoted by (l_F^u, τ_F^u), (l_F^{s1}, τ_F^{s1}), and (l_F^{s2}, τ_F^{s2}), respectively. The corresponding rents are denoted by r_F^u, r_F^{s1}, and r_F^{s2}, respectively. The responses and revenues all depend on l_L and τ_L. For

---

[32] We can ignore Case 1 since setting the same standard level and tax as L is never a best response for F except if $(l_L, \tau_L) = (0, 0)$ and $k \leq 1$, which is treated below. We can also ignore Case 5 since choosing a response that does not attract any firm is never a best response for F.

notational ease, we will use (l_F^*, τ_F^*) to denote the response that maximizes $r_F(l_F, \tau_F; l_L, \tau_L)$ over $\{(l_F^u, \tau_F^u), (l_F^{s1}, \tau_F^{s1}), (l_F^{s2}, \tau_F^{s2})\}$.

There are two issues that can arise when solving for the developing country's best response to (l_L, τ_L): First, a best response might not exist; second, a best response might not be unique. The existence of a best response to (l_L, τ_L) is not guaranteed because an optimal undercutting strategy does not exist. The reason is that the rent function does not have a well-defined maximum on the set of undercutting strategies. That is, for each undercutting strategy with $\tau_F = \tau_L - k|l_L - l_F| - \varepsilon$ where $\varepsilon > 0$, we can find a slightly higher tax (i.e., a smaller ε) that still undercuts L's strategy. Because such a tax would yield higher rents, the optimal undercutting strategy is not well defined.[33]

In our model this difficulty can be resolved in a straightforward way. Since L, whenever possible, avoids strategies that induce F to undercut (i.e., undercutting happens only off the equilibrium path), we can solve our model without determining the undercutting strategy. To see that, let $r_F^s(l_L, \tau_L)$ be F's rent from an optimal sharing strategy, and, given $\varepsilon > 0$, let $r_F^u(l_L, \tau_L; \varepsilon)$ be F's rent from undercutting where $\tau_F = \tau_L - k|l_L - l_F| - \varepsilon$. Let $r_F^u(l_L, \tau_L) = \lim_{\varepsilon \to 0} r_F^u(l_L, \tau_L; \varepsilon)$. Note that, by choosing ε sufficiently small, F can obtain a rent arbitrarily close to $r_F^u(l_L, \tau_L)$, while it still holds that $r_F^u(l_L, \tau_L; \varepsilon) < r_F^u(l_L, \tau_L)$ no matter how small is ε. By solving $r_F^s(l_L, \tau_L) = r_F^u(l_L, \tau_L)$ we obtain $\hat{\tau}_L(l_L)$, the sharing tax limit, for which

if $\tau_L \leq \hat{\tau}_L(l_L)$ then for all $\varepsilon > 0$, it holds that $r_F^s(l_L, \tau_L) > r_F^u(l_L, \tau_L; \varepsilon)$, and

if $\tau_L > \hat{\tau}_L(l_L)$ then there exists an $\varepsilon > 0$ such that $r_F^s(l_L, \tau_L) < r_F^u(l_L, \tau_L; \varepsilon)$.

Thus, if L's tax is higher than the sharing tax limit, then F can find an ε sufficiently small to make undercutting rents higher than the rents earned by sharing. If L sets its tax no higher than the sharing tax limit, sharing yields higher rents for F than undercutting, no matter how small is ε. We assume that F undercuts (by some ε) whenever $\tau_L > \hat{\tau}_L(l_L)$ and shares otherwise.

To deal with the case of F having multiple best responses, recall that by A3 a government with multiple best responses chooses the best response that maximizes its share of firms and if in addition $(l_L, \tau_L) = (0, 0)$, then F sets $\tau_F = 0$. We invoke this assumption in only two

---

[33]The literature on entry deterrence through pricing strategy has also had to broach the issue of what constitutes a best response when payoff functions defined by the game are discontinuous and might not have a well defined maximum. This issue carries over to the present setting.

situations. First, if $k < 1$ and $(l_L, \tau_L) = (0, 0)$, there is no response that yields positive rents for F, so $(l_F, \tau_F) = (0, 0)$.[34] Second, if $k = 1$ then for any (l_L, τ_L) Government F has a best response at each standard level l_F. Since standard mismatch and taxes are equally costly for firms, lowering the standard and the tax by the same amount attracts the same share of firms as before while leaving rents per firm the same. For this case, our assumption implies that $l_F = 0$.

Proposition 1 summarizes the best response of Government F.

Proposition 1 (*The developing country's best response*)

a. *If the marginal cost of standard mismatch for firms is low $(k \leq \frac{1}{3})$, Government F's best response to any of Government L's feasible strategies is to set zero standard, and to set an undercutting tax if $\tau_L > 0$ and to set $\tau_F = 0$ if $\tau_L = 0$. Specifically, if $\tau_L > 0$ then $(l_F^*, \tau_F^*) = (0, \tau_F^u(l_L, \tau_L))$, and if $\tau_L = 0$ then $(l_F^*, \tau_F^*) = (0, 0)$.*

b. *If the marginal cost of standard mismatch is at an intermediate level $(\frac{1}{3} < k \leq 1)$ there exists, for each standard level set by L, a corresponding sharing tax limit. If L sets its tax above (equal to or below) the sharing tax limit, then Government F's best response is to set no standard and to set the corresponding optimal undercutting tax (optimal sharing tax). Specifically, for each l_L there exists a sharing tax limit, $\hat{\tau}_L(l_L)$, such that if $\tau_L > \hat{\tau}_L(l_L)$ then $(l_F^*, \tau_F^*) = (0, \tau_F^u(l_L, \tau_L))$ and if $\tau_L \leq \hat{\tau}_L(l_L)$ then $(l_F^*, \tau_F^*) = (0, \tau_F^{s1}(l_L, \tau_L))$. F's optimal sharing tax is given by $\tau_F^{s1}(l_L, \tau_L) = \frac{1}{2}\tau_L + \frac{k}{2}l_L$.*

c. *If the marginal cost of standard mismatch is high $(k > 1)$ there exists, for each standard level set by L, a corresponding sharing tax limit. If L sets its tax above (equal to or below) the sharing tax limit, then Government F's best response is to set the optimal undercutting tax while setting the same standard level as L (set the optimal sharing tax and set either a lower or higher standard than L). Specifically, for each l_L there exists a sharing tax limit, $\hat{\tau}_L(l_L)$, such that if $\tau_L > \hat{\tau}_L(l_L)$ then $(l_F^*, \tau_F^*) = (l_L, \tau_F^u(l_L, \tau_L))$ and if $\tau_L \leq \hat{\tau}_L(l_L)$ then $(l_F^*, \tau_F^*) \in \{(l_F^{s1}, \tau_F^{s1}(l_L, \tau_L)), (l_F^{s2}, \tau_F^{s2}(l_L, \tau_L))\}$.*

If the marginal cost of standard mismatch is low or at an intermediate level $(k \leq 1)$, it does not pay for Government F to compete in the standard at all. Thus $l_F^* = 0$ in parts (a) and (b). However, if the marginal cost of standard mismatch is high $(k > 1)$, F has an

---

[34]Note that any other feasible strategy for F would induce all firms to locate in L.

incentive to set a positive standard level. Moreover, the cheapest way to attract *all* firms is to set exactly the same level of standard as L. In this way F does not need to compensate any of the firms for a higher standard mismatch. The optimal sharing I and sharing II taxes for case (c) are both boundary solutions. Government F sets the highest tax that still attracts some firms to its country (the firms in the intervals $[0, l_F]$ and $[l_F, 1]$, respectively).

For small k undercutting dominates sharing (Part (a)). All firms can be attracted without having to set the tax much below L's tax. For $\tau_L = 0$ there is no strategy for F that yields a positive rent, meaning that F is indifferent among all feasible strategies. Thus, (by assumption) F sets a zero standard and sets $\tau_F = 0$.

To see why in Part (b) of Proposition 1 F shares if L's tax does not exceed $\hat{\tau}_L(l_L)$ but undercuts otherwise, suppose that, for some standard and tax levels, L and F are sharing firms.. When raising its own tax, F has to consider a 'tax level effect' - F will earn more rent per firm - and a 'tax base effect' - fewer firms will locate in F. Since for $k \leq 1$ firms' location decisions are relatively elastic with respect to taxes (recall that $\hat{s} = \frac{\tau_L - \tau_F}{2k} + \frac{l_L + l_F}{2}$) the tax base effect dominates the tax level effect and F increases its tax less than L does ($\tau_F^{s_1} = \frac{1}{2}\tau_L + \frac{k}{2}l_L$). Thus the share of firms locating in F increases. The more L raises its tax, the more firms it will lose to F. Eventually, all firms with $s \leq l_L$ locate in F. At this point, an undercutting strategy because yields higher rents for F because it has to lower τ_F only marginally to increase its share by $(1 - l_L)$.

Similarly, there exists a sharing tax limit in part (c). Notice that the proposition only states that F shares firms up to that tax level, but not whether it does so by setting a lower or higher standard than L. It is possible (as shown in the proof) to identify two subsets of S_L so that F chooses $(l_F^{s_1}, \tau_F^{s_1})$ or $(l_F^{s_2}, \tau_F^{s_2})$ if (l_L, τ_L) is in the first or second of the subsets, respectively. Intuitively, if L sets a relatively low standard level then sharing with a higher standard level tends to yield higher rents for F; if L sets a relatively high standard level then sharing with a lower standard level yields higher rents for F. In particular, if $l_L \geq \frac{1}{2}$ then setting a higher standard and sharing is never a best response for F. Instead of setting a standard that exceeds L's standard by x, i.e. $l_F = l_L + x$, and some tax τ_F, Government F can set $l_F = l_L - x$ without changing the tax. Doing so increases F's rent per firm without reducing its share of firms.

A.3.2 The developed country's best response

Government L takes Government F's subgame perfect strategy $f^*(l_L, \tau_L)$ as given and maximizes its rent function over S_L. Just like Government F's rent function, L's rent function evaluated at f^* is not continuous. For example, discontinuities arise at the sharing tax limit $\hat{\tau}_L(l_L)$. But given f^*, we can safely exclude from the set of candidates for best response all strategies with $\tau_L > \hat{\tau}_L(l_L)$ (except $(l_L, \tau_L) = (0, 0)$), because such strategies induce F to undercut and hence leave L with zero rents while a tax that induces F to share firms yields positive rents. Thus, we can formulate the following problem for L.

Government L's optimization problem.

$$\max_{\{l_L, \tau_L\}} r_L\left(l_L, \tau_L, f^*(l_L, \tau_L)\right)$$
$$\text{s.t.}$$

$$l_L \in [0, 1]$$
$$\tau_L \in [l_L, \hat{\tau}_L(l_L)]$$

Figure 8 depicts L's rent function, $r_L\left(l_L, \tau_L, f^*(l_L, \tau_L)\right)$, depending on τ_L and fixing some standard level l_L. Rents are zero when L sets $\tau_L = l_L$, but then increase at low levels of τ_L when F is willing to share firms. Rents jump to zero at $\tau_L = \hat{\tau}_L(l_L)$.
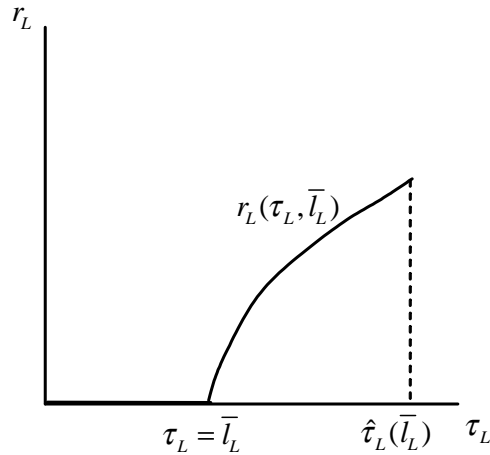

Figure 8

Proposition 2 summarizes L's best response to f^*. We use $\hat{l}_L$ to denote the critical standard level so that the sharing tax limit is at least as large as the cost to cover the standard if and only if $l_L \geq \hat{l}_L$. See Figure 6 for an illustration.

Proposition 2.[35] *(The developed country's best response to f^*).*

a. *If the marginal cost of standard mismatch is low $(k \leq \frac{1}{3})$, Government L's best response to f^* is to set no standard and set zero tax. Specifically, $(l_L^*, \tau_L^*) = (0, 0)$.*

b. *If the marginal cost of standard mismatch is at an intermediate level $(\frac{1}{3} < k \leq 1)$, Government L's best response to f^* is to set a standard strictly larger than $\hat{l}_L$ and set its tax at the sharing tax limit, $\hat{\tau}_L(l_L)$, the highest tax that induces F to share firms. This tax is higher than the tax set by F. As k is increased, standard provision by L decreases from $l_L^* \lesssim 1$ to $l_L^* \simeq \frac{22}{25}$, and rents per firm increase.*

c. *If the marginal cost of standard mismatch is high $(k > 1)$, Government L's best response to f^* is to set a standard of $l_L^* = \frac{8}{9}$ and to set $\tau_L = \hat{\tau}_L^1(l_L)$, the highest tax that induces F to share firms and which exceeds costs by at least a factor of 2. Specifically, $(l_L^*, \tau_L^*) = \left(\frac{8}{9}, \frac{4}{3} + \frac{4}{9}k\right)$.*

If $k \leq \frac{1}{3}$, Government F chooses an undercutting strategy for each tax that exceeds the cost of the standard (Proposition 1). Thus each of L's strategies yields zero rents and, by A3, Government L picks $(l_L^*, \tau_L^*) = (0, 0)$. Rents for both governments are zero.

For $\frac{1}{3} < k \leq 1$, there exists a (k-dependent) critical standard level $\hat{l}_L$ such that $\hat{\tau}_L(l_L) = l_L$ for all $l_L \leq \hat{l}_L$ and $\hat{\tau}_L(l_L) > l_L$ for all $l_L > \hat{l}_L$. The lower the standard level that L sets, the larger its share of firms and the greater the incentive for F to switch to an undercutting strategy because the switch induces *all* firms located in L to move to F. Therefore, so that it does not induce F to undercut, L puts itself into a situation in which it attracts only a relatively small share of firms by setting a high standard level. Because L's rents are increasing in τ_L up to $\hat{\tau}_L(l_L)$ (see Figure 8), L chooses $\hat{\tau}_L(l_L)$.

When $k > 1$, it is better for L to let F be the country that sets a low standard and earn high sharing rents, because that means that F accommodates higher taxes by L without undercutting. For example, suppose that L chooses $l_L = \frac{1}{9}$ and sets a tax $\tau_L = \hat{\tau}_L\left(\frac{1}{9}\right)$ instead of its actual equilibrium choice $l_L^* = \frac{8}{9}$ and $\tau_L^* = \hat{\tau}_L\left(\frac{8}{9}\right)$. Government F's best response would be to set $l_F = \frac{1}{3}$ instead of $l_F^* = \frac{2}{3}$. As with the actual equilibrium strategies, L attracts one third of the firms. But the tax L is able to set, $\hat{\tau}_L\left(\frac{1}{9}\right)$, is so much lower than $\hat{\tau}_L\left(\frac{8}{9}\right)$ that rents per firm are only $\frac{4}{9}k - \frac{4}{9}$ compared to $\frac{4}{9}k + \frac{4}{9}$ with the equilibrium strategy.

---

[35]Proposition 2 is restated below with the exact expressions for the optimal strategies.

In order to obtain the ability to set a higher tax without losing firms, L accepts that it has to set a costlier standard level.

Propositions 1 and 2 can be used to solve for the mutual best responses of the strategies of F and L, thus yielding the subgame perfect Nash equilibrium in pure strategies presented in Theorem 1.

A.3.3 Proofs

The proof of Proposition 1 uses a sequence of auxiliary results, which are stated and proven separately in the following Lemmas.

Lemma 1.

1. If $k < 1$, undercutting is feasible if and only if $(l_L, \tau_L) \in S_L \setminus \{(0,0)\}$. Undercutting with $l_F > 0$ is never a best response.

2. If $k = 1$, undercutting is feasible if and only if $\tau_L > l_L$. For every undercutting strategy with $l_F > 0$, there exists an undercutting strategy with $l_F = 0$ that yields the same rent for F.

3. If $k > 1$, undercutting is feasible if and only if $(l_L, \tau_L) \in S_L$ such that $\tau_L > l_L$. Undercutting with $l_F \neq l_L$ is never a best response.

Proof.

1. We will first show that undercutting I is non-empty if $k < 1$ and $(l_L, \tau_L) \neq (0,0)$. Let (l_L, τ_L) be any strategy in $S_L \setminus \{(0,0)\}$. Set $l_F = 0$. Then for small enough ε, l_F together with the tax $\tau_F = \tau_L - k l_L - \varepsilon \geq 0$ is a feasible undercutting strategy. If $(l_L, \tau_L) = (0,0)$ undercutting is not feasible, because it requires to set a tax strictly below zero, which is not feasible. Next, we show that $(l_F^u, \tau_u^u) = (0, \tau_L - k l_L - \varepsilon)$ for some $\varepsilon > 0$. Take any undercutting strategy with $l_F > 0$ and a corresponding undercutting tax $\tau_F = \tau_L - k |l_L - l_F| - \varepsilon$. Using the same ε to undercut, the strategy $l_F' = 0$ with undercutting tax $\tau_F' = \tau_L - k l_L - \varepsilon$ is feasible (i.e., $\tau_L - k l_L - \varepsilon > 0$) and yields more rents per firm because it saves costs of l_F per firm and reduces revenue per firm by at most $k l_F'$. Thus, undercutting with $l_F > 0$ is never a best response.

2. If $k = 1$, it is obvious that undercutting is feasible if $\tau_L > l_L$: simply let $l_F = l_L$ and choose ε such that $\tau_F = \tau_L - \varepsilon \geq l_L$. To see that the reverse implication holds, suppose

that $\tau_L = l_L$. In this case F cannot find a strategy so that the firm $s = l_L$ prefers the tax and the standard level offered in F to the ones offered in L because if $l_F \neq l_L$, F will have to compensate $s = l_L$ for more than its standard mismatch meaning that F's tax would have to undercut L's tax by more than $|l_L - l_F|$ which is not feasible. Next, fix (l_L, τ_L) and let (l_F, τ_F) be a feasible undercutting strategy with $l_F > 0$. The strategy (l'_F, τ'_F) with $l'_F = 0$ and $\tau'_F = \tau_F - l_F$ yields the same rent per firm as (l_F, τ_F). Moreover, the fact that every firm preferred (l_F, τ_F) to (l_L, τ_L) implies that every firm also prefers (l'_F, τ'_F) to (l_L, τ_L) (all firms $s < l_F$ strictly prefer (l'_F, τ'_F) to (l_F, τ_F) and all other firms are indifferent).

3. Suppose that $k > 1$. If $\tau_L > l_L$, undercutting is obviously feasible. If $\tau_L = l_L$, undercutting is not feasible. Take any $l_F \in [0, 1]$. We have

$$\tau_F = \tau_L - k\,|l_L - l_F| - \varepsilon < l_F.$$

Next, we show that if undercutting is feasible then $(l_F^u, \tau_u^u) = (l_L, \tau_L - \varepsilon)$ for some $\varepsilon > 0$. Assume that $\tau_L > l_L$. Take any undercutting strategy with $l_F \neq l_L$ and a corresponding undercutting tax $\tau_F = \tau_L - k\,|l_L - l_F| - \varepsilon$. Using the same ε to undercut, the strategy $l'_F = l_L$ with undercutting tax $\tau'_F = \tau_L - \varepsilon$ is feasible. Comparing rents per firm if $l_F < l_L$, we get that

$$
\begin{aligned}
\tau_L - \varepsilon - l_L \;\; &> \;\; \tau_L - k(l_L - l_F) - \varepsilon - l_F \Longleftrightarrow \\
l_L\,(k-1) \;\; &> \;\; l_F\,(k-1)
\end{aligned}
$$

which is true for $k > 1$. If $l_F > l_L$, we get that

$$
\begin{aligned}
\tau_L - \varepsilon - l_L \;\; &> \;\; \tau_L - k(l_F - l_L) - \varepsilon - l_F \Longleftrightarrow \\
l_L\,(1-k) \;\; &> \;\; l_F\,(1-k)
\end{aligned}
$$

which is true as well, showing that for any undercutting strategy $l_F \neq l_L$ there exists another undercutting strategy yielding more rents. $\square$

In the following, we will deal with the case $(l_L, \tau_L) \neq (0, 0)$. If $(l_L, \tau_L) = (0, 0)$, by Lemma 1, undercutting is not feasible, and any feasible strategy for F yields zero rents. By assumption, F chooses $(l_F, \tau_F) = (0, 0)$.

Lemma 2a.

1. If $k < 1$, for any (l_L, τ_L), a sharing strategy is optimal among strategies in sharing I and sharing II only if $l_F = 0$.

2. If $k = 1$, for any (l_L, τ_L), there exists a best response (l_F, τ_F) to (l_L, τ_L) such that $l_F = 0$.

Proof.

1. Take any sharing strategy (l_F, τ_F) such that $0 < l_F \leq l_L$. Let $(l'_F, \tau'_F) = (0, \tau_F - k l_F)$. This strategy is feasible, attracts the same fraction of firms, and F makes strictly higher rents per firm. Next, take any sharing strategy (l_F, τ_F) such that $l_F > l_L$. The strategy $(l'_F, \tau'_F) = (l_F - \varepsilon, \tau_F - \varepsilon)$ such that $l_F - \varepsilon > l_L$ is feasible for small enough ε and yields strictly higher rents for jurisdiction F.

2. We will proof the statement by showing that for any sharing strategy with $l_F > 0$ there exists a sharing strategy with $l_F = 0$ that yields the same rent. Fix (l_L, τ_L) and let $(l_F, \tau_F) \in sharing\ I$. Consider the strategy (l'_F, τ'_F) with $l'_F = 0$ and $\tau'_F = \tau_F - l_F$. This strategy yields the same rent per firm, so it suffices to show that the same firms locate in F under $((l_L, \tau_L), (l_F, \tau_F))$ as under $((l_L, \tau_L), (l_F, \tau_F))$. Suppose s (weakly) preferred F to L under $((l_L, \tau_L), (l_F, \tau_F))$. If $s \leq l_F$, then

$$|l'_F - s| + \tau'_F = s + \tau_F - l_F \leq s - l_F + \tau_F,$$

so s (at least weakly) prefers (l'_F, τ'_F) to (l_F, τ_F), implying that s also prefers (l'_F, τ'_F) to (l_L, τ_L). If $s > l_F$, then

$$|l'_F - s| + \tau'_F = s - l_F + \tau_F,$$

so s is indifferent between (l'_F, τ'_F) to (l_F, τ_F). The proof for $(l_F, \tau_F) \in sharing\ II$ is analogous. $\square$

Lemma 2b.

1. If $k < 1$, the unique rent maximizing sharing strategy for F is

$$(l^s_F, \tau^s_F) = \begin{cases} \left(0, \frac{1}{2}\tau_L + \frac{k}{2}l_L\right) & \text{if} \quad \tau_L \leq 3kl_L \\ (0, \tau_L - kl_L) & \tau_L > 3kl_L \end{cases}$$

33

2. If $k = 1$, the sharing strategy

$$(l_F^s, \tau_F^s) = \begin{cases} (0, \frac{1}{2}\tau_L + \frac{1}{2}l_L) & \text{if} \quad \begin{matrix} \tau_L \leq 3l_L \\ \tau_L > 3l_L \end{matrix} \\ (0, \tau_L - l_L) \end{cases}$$

maximizes rents.

Proof.

1. From Lemmas 1 and 2a we know that $l_F = 0$ at any best response of F. We will derive the optimal sharing tax and show that there always exists an ε such that undercutting yields more rents. Given (τ_L, l_L), government F faces the following optimization problem for sharing,

$$\max_{\tau_F} \left\{ \tau_F \left(\frac{\tau_L - \tau_F}{2k} + \frac{l_L}{2} \right) \right\} \qquad ((*))$$

$$\text{s.t. } \hat{s}(l_L, \tau_L, 0, \tau_F) \in [0, l_L]$$

$$\tau_F \geq 0$$

We will ignore the constraints for the moment. The revenue function is strictly concave in τ_F, so our solution will be unique and we only need to consider first-order conditions

$$\frac{\partial}{\partial \tau_F} \left(\tau_F \left(\frac{\tau_L - \tau_F}{2k} + \frac{l_L}{2} \right) \right) = \frac{1}{2}l_L - \frac{1}{2k}\tau_F^* + \frac{1}{2k}(\tau_L - \tau_F^*) = 0$$

$$\iff \tau_F^* = \frac{1}{2}\tau_L + \frac{k}{2}l_L$$

Obviously, $\tau_F^* \geq 0$, so we only need to verify whether $\hat{s}(l_L, \tau_L, 0, \tau_F^*) \in [0, l_L]$. We have

$$\hat{s}(l_L, \tau_L, 0, \tau_F^*) = \frac{1}{4k}\tau_L + \frac{l_L}{4},$$

which is strictly larger than zero. But

$$\hat{s}(l_L, \tau_L, 0, \tau_F^*) \leq l_L \iff \tau_L \leq 3kl_L$$

So, if $\tau_L \leq 3kl_L$, one of the constraints binds. Strategies with $\tau_F = 0$ or $\hat{s}(\tau_L, l_L, \tau_F, 0) = 0$ yield zero rents. A strategy with $\hat{s}(l_L, \tau_L, 0, \tau_F) = l_L$, i.e. $\tau_F = \tau_L - kl_L$, yields $r_F(l_L, \tau_L) = (\tau_L - kl_L)l_L > 0$ if $l_L > 0$ (if $l_L = 0$, then $\tau_L \leq 3kl_L$ implies $\tau_L = 0$, and we do not consider such strategies here).

2. The proof is analogous to the proof or Part 1, except that we do not get uniqueness. $\square$

Lemma 3. *If $k \leq \frac{1}{3}$, the rent maximizing undercutting strategy yields higher rents than the rent maximizing sharing strategy for all (l_L, τ_L) such that $\tau_L > l_L$.*

Proof. For $k \leq \frac{1}{3}$, we have $\tau_L > 3kl_L$ for all strategies with $\tau_L > l_L$. so by Lemma 2b the optimal sharing strategy is $(l_F^s . \tau_F^s) = (0, \tau_L - kl_L)$. The corresponding rents are $r_F^s (l_L, \tau_L) = (\tau_L - kl_L) l_L$ (note that our assumptions imply that, for τ_F such that $\hat{s}(l_L, \tau_L, 0, \tau_F) = l_L$, all firms $s \geq l_L$ locate in L) Comparing this to the rents from undercutting shows that, for ε small enough (notice ε depends on l_L), undercutting rents are better. If $l_L < 1$,

$$r_F^v (\tau_L, l_L) = (\tau_L - kl_L - \varepsilon) > (\tau_L - kl_L) l_L = r_F^{s1} (\tau_L, l_L)$$

for ε sufficiently small. If $l_L = 1$, the optimal sharing strategy is in fact an undercutting strategy (it attracts all firms but a set of firms of measure zero). $\square$

Lemma 4. *Let $\frac{1}{3} < k \leq 1$. For each $l_L \in [0, 1]$, there exists a $\hat{\tau}_L (l_L; k)$ such that $r_F^s > r_F^u$ for all $\tau_L \leq \hat{\tau}_L (l_L; k)$ and $r_F^u > r_F^s$ for all $\tau_L > \hat{\tau}_L (l_L; k)$.*

Proof. By Lemma 2b, if $\tau_L > 3kl_L$, then the optimal sharing strategy is not interior, and the proof of Lemma 3 shows that undercutting is better than sharing.[36] It only remains to consider the case $\tau_L \leq 3kl_L$. Optimal sharing revenues are given by

$$r_F^s (\tau_L, l_L) = \left(\frac{1}{2}\tau_L + \frac{k}{2}l_L \right) \left(\frac{1}{4k}\tau_L + \frac{l_L}{4} \right)$$

For $l_L < 1$, sharing yields more rents than undercutting if and only if

$$\left(\frac{1}{2}\tau_L + \frac{k}{2}l_L \right) \left(\frac{1}{4k}\tau_L + \frac{l_L}{4} \right) > (\tau_L - kl_L - \varepsilon)$$

We now set $\varepsilon = 0$ and solve for the tax at which both sides are equal. This tax will be the highest tax that L can set so that F does not undercut. No matter how small F sets ε, the right hand side will be smaller then the left hand side at this tax. On the other hand, for a tax that is larger than the tax at which both sides are equal, F can find an ε sufficiently small that undercutting yields higher rents than sharing. We solve

$$\left(\frac{1}{2}\tau_L + \frac{k}{2}l_L \right) \left(\frac{1}{4k}\tau_L + \frac{l_L}{4} \right) = (\tau_L - kl_L) \Longleftrightarrow$$

$$kl_L - \tau_L + \frac{1}{8k} \left(2kl_L\tau_L + \tau_L^2 + k^2 l_L^2 \right) = 0$$

---

[36] Notice that at $l_L = 0$, this always holds so that undercutting is always better.

The left hand side expression is a quadratic function of τ_L. Solving the equation yields two solutions, which we denote by $\tilde{\tau}_L^1(l_L, k)$ and $\tilde{\tau}_L^2(l_L, k)$. They are given by

$$\tilde{\tau}_L^1(l_L, k) = k\left(4 - 4\sqrt{1 - l_L} - l_L\right)$$
$$\tilde{\tau}_L^2(l_L, k) = k\left(4 + 4\sqrt{1 - l_L} - l_L\right)$$

Notice that the factor in front of τ_L^2 is positive. Sharing revenues are therefore larger than undercutting revenues for $\tau_L \leq \tilde{\tau}_L^1(l_L, k)$. Because $\tilde{\tau}_L^2(l_L, k) > 3kl_L > \tilde{\tau}_L^1(l_L, k)$ undercutting revenues are higher for all $\tau_L > \tilde{\tau}_L^1(l_L, k)$. It can be verified that $\tilde{\tau}_L^1(l_L, k) \leq l_L$ for $l_L \leq 8k\frac{1-k}{(1+k)^2}$, and $\tilde{\tau}_L^1(l_L, k) > l_L$ for $l_L > 8k\frac{1-k}{(1+k)^2}$ (we omit the derivation). Therefore the critical tax beyond which F will undercut is given by

$$\hat{\tau}_L(l_L, k) = \begin{cases} l_L & \text{if } l_L \leq 8k\frac{1-k}{(1+k)^2} \\ k\left(4 - 4\sqrt{1 - l_L} - l_L\right) & \text{otherwise} \end{cases}$$

See also Figure 5 in Section 4. $\square$

Lemma 5. *Let $k > 1$. The strategy that maximizes $r_F(l_F, \tau_F; l_L, \tau_L)$ over sharing I is given by $l_F^{s1} = \frac{kl_L + \tau_L}{2(k+1)}$ and $\tau_F^{s1} = \frac{(\tau_L + kl_L)(k+2)}{2(k+1)}$. The strategy that maximizes $r_F(l_F, \tau_F; l_L, \tau_L)$ over sharing II is given by $l_F^{s2} = \frac{k(1+l_L) - (1+\tau_L)}{2(k-1)}$ and $\tau_F^{s2} = \frac{(\tau_L - kl_L)(k-2) + k(k-1)}{2(k-1)}$.*

Proof. We start with deriving the optimal sharing strategy over sharing I. Government F's problem is

$$\max_{(l_F, \tau_F)} \left\{(\tau_F - l_F)\left(\frac{\tau_L - \tau_F}{2k} + \frac{l_L + l_F}{2}\right)\right\}$$
$$\text{s.t.}$$

$$\tau_F \geq l_F$$
$$l_F \in [0, l_L)$$
$$\tau_F \in [\tau_L - k(l_L - l_F), \tau_L + k(l_L - l_F)]$$

Without doing the calculus, we will reduce the optimization problem by first showing that a necessary condition for (l_F, τ_F) being a solution to the problem is that $\tau_F = \tau_L + k(l_L - l_F)$, i.e., given some l_F, Government F will set the highest tax that possibly attracts some firms to its jurisdiction. Take any strategy (l_F, τ_F) with $\tau_F < \tau_L + k(l_L - l_F)$ (notice that these are the strategies that are not at the upper bound of the sharing I set, see also Figure 2). Compare this strategy to another strategy (l_F', τ_F') with $l_F' = l_F + \delta$ and $\tau_F' = \tau_F + \delta$, where

36

$\delta > 0$. For a sufficiently small δ, the pair (l'_F, τ'_F) is in sharing I. This strategy yields the same rents per firm but attracts more firms to F because

$$
\begin{aligned}
\hat{s}\left(l_F, \tau_F, l_L, \tau_L\right) &< \hat{s}\left(l'_F, \tau'_F, l_L, \tau_L\right) \Longleftrightarrow \\
\frac{\tau_L - \tau_F}{2k} + \frac{l_L + l_F}{2} &< \frac{\tau_L - \tau'_F}{2k} + \frac{l_L + l'_F}{2} \Longleftrightarrow \\
-\tau_F + k l_F &< -\tau_F - \delta + k\left(l_F + \delta\right) \Longleftrightarrow \\
1 &< k
\end{aligned}
$$

for $\delta > 0$.

Therefore, we can reduce F's problem to

$$
\max_{l_F} \left\{ \left(\tau_L + k\left(l_L - l_F\right) - l_F\right) l_F \right\}
$$

s.t.

$$
l_F \in [0, l_L)
$$

The objective function is strictly concave in l_F, so second order conditions will be satisfied, and the maximizer is unique. Ignoring the constraint for the moment and solving for an interior solution yields

$$
\begin{aligned}
\frac{\partial}{\partial l_F}\left(\left(\tau_L + k\left(l_L - l_F\right) - l_F\right) l_F\right) &= -2l_F k + l_L k - 2l_F + \tau_L = 0 \Longleftrightarrow \\
l_F^{s1} &= \frac{k l_L + \tau_L}{2\left(k + 1\right)}
\end{aligned}
$$

Obviously, $l_F^{s1} \geq 0$. But

$$
\begin{aligned}
l_F^{s1} &\leq l_L \Longleftrightarrow \\
k l_L + \tau_L &\leq l_L 2\left(k + 1\right) \Longleftrightarrow \\
\tau_L &\leq l_L\left(k + 2\right)
\end{aligned}
$$

For higher τ_L, sharing with less standard is not the optimal strategy. At the boundary solution $l_F^{s1} = l_L$ undercutting yields more than sharing. The corresponding tax F sets would be $\tau_F = \tau_L = l_L\left(k + 2\right)$. By assumption, it would attract half of the firms and therefore

$$
\begin{aligned}
r_F^{s1}\left(l_L, \tau_L\right) &= \left(\tau_L - l_L\right)\frac{1}{2} \\
&< \left(\tau_L - \varepsilon - l_L\right) \\
&= r_F^u\left(l_L, \tau_L\right)
\end{aligned}
$$

37

for sufficiently small ε. Therefore undercutting is better than the optimal sharing I strategy if $\tau_L > l_L(k+2)$.

The strategy that maximizes r_F over sharing II can be derived analogously. We omit this derivation here, but note that

$$
\begin{aligned}
l_F^{s2} &\geq\; l_L \Longleftrightarrow \\
1 - l_L k - k + \tau_L &\leq\; l_L(2 - 2k) \Longleftrightarrow \\
\tau_L &\leq\; -1 + 2l_L + k - kl_L.
\end{aligned}
$$

So again, we get a bound for τ_L so that the optimal undercutting strategy yields higher rents than the strategy that maximizes r_F over sharing II if τ_L is larger than this bound. $\square$

Lemma 6. Let $k > 1$. For each $l_L \in [0,1]$, there exists a $\hat\tau_L^1(l_L; k)$ such that $r_F^{s1} > r_F^u$ for all $\tau_L \leq \hat\tau_L^1(l_L; k)$ and $r_F^u > r_F^{s1}$ for all $\tau_L > \hat\tau_L^1(l_L; k)$, and a $\hat\tau_L^2(l_L; k)$ such that $r_F^{s2} > r_F^u$ for all $\tau_L \leq \hat\tau_L^2(l_L; k)$ and $r_F^u > r_F^{s2}$ for all $\tau_L > \hat\tau_L^2(l_L; k)$.

Proof: We first derive $\hat\tau_L^1(l_L; k)$. Suppose $\tau_L \leq (1 - l_L)(k+2)$ (recall from the proof of Lemma 5 that this was the upper bound for τ_L, so that the constraint $l_F \leq 1 - l_L$ was not binding). We will derive $\hat\tau_L^1(l_L; k)$ and then verify that it is indeed not larger than this bound, so that undercutting is better than the optimal sharing I strategy for all $\tau_L > \hat\tau_L^1(l_L; k)$. For given (l_L, τ_L) rents from the optimal sharing 1 strategy are given by

$$
r_F^{s1}(l_L, \tau_L) = \frac{1}{4}(k+1)^{-1}(l_L k + \tau_L)^2
$$

The derivation of $\hat\tau_L^1(l_L; k)$ is analogous to the derivation of $\hat\tau_L(l_L; k)$ in the proof of Lemma 4, so we provide less detail. Let $\varepsilon = 0$, and set the difference of this rent and undercutting rents equal to zero. We can solve for the highest tax of government L depending on l_L such that F prefers sharing to undercutting[37]:

$$
\begin{aligned}
r_F^{s1}(l_L, \tau_L) - r_F^u(l_L, \tau_L) &=\; \frac{1}{4}(k+1)^{-1}(l_L k + \tau_L)^2 - (\tau_L - l_L) = 0 \\
&\Longleftrightarrow\; \hat\tau_L^1(l_L; k) = 2 - l_L k + 2k - 2(k+1)\sqrt{1 - l_L}
\end{aligned}
$$

It can be verified that $l_L \leq \hat\tau_L^1(l_L) \leq l_L(k+2)$, and therefore, for all $\tau_L \leq \hat\tau_L^1(l_L)$ Government F prefers the strategy with l_F^{s1} and the highest sharing tax to the optimal undercutting strategy, and prefers undercutting otherwise.

---

[37] As in the proof of Lemma 4, we obtain two solutions but the second one will be larger than $l_L(2+k)$

The derivation for $\hat{\tau}_L^2(l_L)$ exactly parallels the one for $\hat{\tau}_L^1(l_L)$. The rent difference for $\varepsilon = 0$ can be solved to obtain $\hat{\tau}_L^2(l_L)$, which can also be verified to be no smaller than l_L

$$r_F^{s2}(l_L, \tau_L) - r_F^u(\tau_L, l_L) = \frac{1}{4}(k-1)^{-1}(k - l_L k - 1 + \tau_L)^2 - (\tau_L - l_L) = 0$$

$$\iff \hat{\tau}_L^2(l_L) = -z + l_L k + k - 2(k-1)\sqrt{l_L}$$

Again we can verify that $l_L \leq \hat{\tau}_L^2(l_L) \leq -1 + 2l_L + k - kl_L$, and therefore, for all $\tau_L \leq \hat{\tau}_L^2(l_L)$ Government F prefers the strategy with l_F^{s2} and the highest sharing tax to the optimal undercutting strategy, and prefers undercutting otherwise. $\square$

Proof of Proposition 1.

Part a follows from Lemmas 1,2a,2b, and 3.

Part b follows from Lemmas 1, 2a, 2b, and 4.

Part c follows from Lemmas 5 and 6. $\square$

Proposition 2.[38] *(The developed country's best response to f^*).*

a. If $k \leq \frac{1}{3}$, then $(l_L^*, \tau_L^*) = (0, 0)$.

b. If $\frac{1}{3} < k \leq 1$, then $(l_L^*, \tau_L^*) = \left(\begin{array}{c} 1 - \left(\frac{1}{3k+3}\left(4k - \sqrt{3 - 6k + 7k^2}\right)\right)^2, \\ \frac{2k}{9}(k+1)^{-2}\left(k^2 + 6\sqrt{7k^2 - 6k + 3} + 2k\sqrt{7k^2 - 6k + 3} + 15\right) \end{array} \right)$.

c. If $k > 1$, then $(l_L^*, \tau_L^*) = \left(\frac{8}{9}, \frac{4}{3} + \frac{4}{9}k\right)$.

Proof:

Part a: By Proposition 1, $f^*(l_L, \tau_L) = (0, \tau_F^u(l_L, \tau_L))$ for all $(l_L, \tau_L) \in S_L \setminus \{(0,0)\}$. By assumption, $f^*(0,0) = (0,0)$. Therefore $r_L(l_L, \tau_L, f(l_L, \tau_L)) = 0$ for all $(l_L, \tau_L) \in S_L$. Using our assumptions again, we obtain $(l_L^*, \tau_L^*) = (0, 0)$. $\square$

Part b: From Proposition 1, we know that, for each level l_L, Government F is going to locate at $l_F = 0$ and undercut if $\tau_L > \hat{\tau}_L(l_L)$. Such strategies can therefore not be optimal for government L, because it can assure itself of positive rents by setting $l_L = 1$ and $\tau_L \in (1, 3k)$ (by Lemma 4, F would choose a sharing strategy in this case). We can also exclude strategies with $l_L = 0$ as F is going to undercut then for every positive tax. The

[38]Proposition 2 and Theorem 1 are restated here with the exact expressions for the optimal strategies.

reduced optimization problem for L is therefore

$$\max_{(\tau_L, l_L)} \{(\tau_L - l_L)(1 - \hat{s}(\tau_L, l_L, \tau_F^s, 0))\}$$

$$\text{s.t. } l_L \in [0, 1]$$

$$\tau_L \in [l_L, \hat{\tau}_L(l_L)]$$

The objective function is continuous and the feasible set is compact. Hence, there exists a solution to the problem. As previously, we will first ignore the constraints, which yields

$$l_L = 4k(k+1)^{-1} > 1$$

So, an interior solution does not exist. At least one of the four constraints is binding. We can exclude $\tau_L = l_L$ and $l_L = 0$ as both strategies yield zero rents.

Case 1) Suppose $\tau_L = \hat{\tau}_L(l_L)$. We will derive the optimal l_L by considering the two cases, $l_L = 1$ and $l_L \in (0, 1)$, separately and then compare the corresponding rents.

(i) $l_L = 1$

This yields rents of $r_L(\hat{\tau}_L(1), 0) = 0$ (because $\hat{s} = 1$).

(ii) $l_L \in (0, 1)$

The maximization problem is

$$\max_{l_L}$$

$$\left\{ \left(k\left(4 - 4\sqrt{1 - l_L} - l_L\right) - l_L \right) \left(1 - \frac{\frac{1}{2}\left(k\left(4 - 4\sqrt{1 - l_L} - l_L\right)\right) - \frac{k}{2}l_L}{2k} - \frac{l_L}{2} \right) \right\}$$

$$\text{s.t. } l_L \in (0, 1)$$

The solution to which is $l_L^* = 1 - \left(\frac{1}{3k+3}\left(4k - \sqrt{3 - 6k + 7k^2}\right)\right)^2 \in (0, 1)$. It can be verified that at l_L^* indeed $\hat{\tau}(l_L^*, k) > l_L^*$ (i.e., $l_L^* > \hat{l}_L$). We denote the corresponding rents by $r_L^1(\tau_F^*, 0)$. They are given by

$$r_L^1(\tau_F^*, 0) = \left(\frac{2}{27}\right)(k+1)^{-2}\left(6k + k^2 + 2k\sqrt{7k^2 - 6k + 3} - 3\right)\left(4k - \sqrt{7k^2 - 6k + 3}\right) > 0.$$

Case 2) Consider a strategy with $l_L = 1$. Maximizing rents with respect tax yields $\tau_L^* = \frac{3}{2}k + \frac{1}{2}$, which is indeed less than $\hat{\tau}_L(1) = 3k$. We denote the corresponding rents by $r_L^2(\tau_F^*, 0)$. They are given by $r_L^2(\tau_F^*, 0) = \frac{1}{16}k^{-1}(3k - 1)^2$.

It can be verified that the inequality

$$r_L^1(\tau_F^*, 0) > r_L^2(\tau_F^*, 0) \iff$$

$$\left(\frac{2}{27}\right)(k+1)^{-2}\left(6k + k^2 + 2k\sqrt{7k^2 - 6k + 3} - 3\right)\left(4k - \sqrt{7k^2 - 6k + 3}\right) > \frac{1}{16}k^{-1}(3k-1)^2$$

holds. We omit the details.

The corresponding tax for government L is

$$\tau_L^* = \frac{2}{9}(k+1)^{-2}\left(k^2 + 6\sqrt{7k^2 - 6k + 3} + 2k\sqrt{7k^2 - 6k + 3} + 15\right)k.$$

Part c: By Lemma 6, we know that $r_F^u(l_L, \tau_L) > r_F^{s1}(l_L, \tau_L)$ if $\tau_L > \hat{\tau}_L^1(l_L)$ and $r_F^u(l_L, \tau_L) > r_F^{s2}(l_L, \tau_L)$ if $\tau_L > \hat{\tau}_L^2(l_L)$. If follows that $f^*(l_L, \tau_L) = (l_F^u, \tau_F^u)$ if $\tau_L > \max\left[\hat{\tau}_L^1(l_L), \hat{\tau}_L^2(l_L)\right]$. If not we, we need to compare r_F^{s1} with r_F^{s2}. If $r_F^{s1} \geq r_F^{s2}$, the optimal strategy must be $f^*(l_L, \tau_L) = (l_F^{s1}, \tau_F^{s1})$, as stated in Lemma 5. If $r_F^{s2} \geq r_F^{s1}$, the optimal strategy must be $f^*(l_L, \tau_L) = (l_F^{s2}, \tau_F^{s2})$, as stated in Lemma 5.

Turning to L, we take f^* as given and first exclude strategies such that $\tau_L > \max\left[\hat{\tau}_L^1(l_L), \hat{\tau}_L^2(l_L)\right]$ as those yield zero rents, while the strategy $(l_L, \tau_L) = \left(0, \hat{\tau}_L^2(0)\right)$ yields strictly positive rents (for this choice, F's best response is (l_F^{s2}, τ_F^{s2}), L attracts a positive fraction of firms, and $\hat{\tau}_L^2(0) > 0$). From the reduced set of possibly optimal strategies for L, we proceed as follows to determine the rent maximizing strategy. First, we show that for (l_L, τ_L) with $l_L \geq \frac{1}{2}$, (l_F^{s2}, τ_F^{s2}) is not a best response for F. We then separately derive the optimal strategies for Government L under two different assumptions:

1. supposing that $r_F^{s1} \geq r_F^{s2}$ so that, in the second stage, F chooses (l_F^{s1}, τ_F^{s1}) if $\tau_L \leq \hat{\tau}_L^1(l_L)$ and undercuts otherwise.

2. supposing that, if $l_L \in \left[0, \frac{1}{2}\right]$, $r_F^{s2} \geq r_F^{s1}$, so that, in the second stage, F chooses (l_F^{s2}, τ_F^{s2}) if $\tau_L \leq \hat{\tau}_L^2(l_L)$ and $l_L \in \left[0, \frac{1}{2}\right]$, and undercuts otherwise.

We will then show that the optimal strategy under supposition 1 yields more rents than the one under supposition 2, and verify that, under this optimal strategy, government F indeed sets less standard and sets the highest sharing tax.

To see that, if $l_L \geq \frac{1}{2}$, setting more standard and setting the highest sharing tax can never be the best response for government F, observe that any strategy (l_F, τ_F) with $l_F > l_L$ and $\tau_F = \tau_L + k(l_F - l_L)$ is dominated by the strategy (l_F', τ_F') with $l_F' = l_L - (l_F - l_L) = 2l_L - l_F$ and $\tau_F' = \tau_F$.

41

1. Now suppose that $r_F^{s1} \geq r_F^{s2}$.

Under this supposition, government L's problem is

$$\max_{(\tau_L, l_L)} \left\{ (\tau_L - l_L)\left(1 - l_F^{s1}\right) \right\}$$

$$\max_{(\tau_L, l_L)} \left\{ (\tau_L - l_L)\left(1 - \frac{k l_L + \tau_L}{2(k+1)}\right) \right\}$$

$$\text{s.t. } l_L \in [0, 1]$$

$$\tau_L \in \left[l_L, \hat{\tau}_L^1(l_L)\right]$$

Solving for an interior solution yields $\tau_L = 1 + \frac{1}{2}l_L - \frac{1}{2}l_L k + k$. But $1 + \frac{1}{2}l_L - \frac{1}{2}l_L k + k \leq \hat{\tau}_L^1(l_L)$ if and only if $l_L \geq 4\sqrt{3} - 6 \approx 0.9282$. Hence, the tax constraint binds for all $l_L \leq 4\sqrt{3} - 6$. Substituting $\hat{\tau}_L^1(l_L)$ into the objective function, we solve the following problem

$$\max_{l_L} \left\{ \left(\hat{\tau}_L^1(l_L) - l_L\right)\left(1 - \frac{k l_L + \hat{\tau}_L^1(l_L)}{2(k+1)}\right) \right\}$$

$$\text{s.t. } l_L \in \left[0, 4\sqrt{3} - 6\right]$$

Solving this for l_L^* yields two solutions, $l_L^{*1} = \frac{8}{9}$ and $l_L^{*2} = 0$. Checking the second-order condition clarifies that only $l_L^{*1} = \frac{8}{9}$ is a maximizer. For simpler notation we write $l_L^{*1} = l_L^*$. Notice that, indeed, $\frac{8}{9} \leq 4\sqrt{3} - 6$. The corresponding revenues are given by $r_L\left(\hat{\tau}_L^1, l_L^*, \tau_F^{s1}, l_F^{s1}\right) = \frac{4}{27}(k+1)$. We also need to verify whether a strategy with $l_L > 4\sqrt{3} - 6$ and no binding tax constraint yields more revenue. The partial derivative with respect to l_L is always positive, and therefore government L wants to set l_L as high as possible. We only need to check $l_L = 4\sqrt{3} - 6$. It can be verified that this strategy does not yield higher rents. The derivation is omitted.

2. Next, suppose that $r_F^{s2} \geq r_F^{s1}$ for $l_L \in \left[0, \frac{1}{2}\right]$.

Under this supposition, government L's problem is

$$\max_{(\tau_L, l_L)} \left\{ (\tau_L - l_L)\left(1 - l_F^{s2}\right) \right\}$$

$$\max_{(\tau_L, l_L)} \left\{ (\tau_L - l_L)\left(\frac{1 - l_L k - k + \tau_L}{2 - 2k}\right) \right\}$$

$$\text{s.t. } l_L \in \left[0, \frac{1}{2}\right]$$

$$\tau_L \in \left[l_L, \hat{\tau}_L^2(l_L)\right]$$

Solving for an interior solution yields $\tau_L = \frac{1}{2}k + \frac{1}{2}l_L k - \frac{1}{2} + \frac{1}{2}l_L$. But $\frac{1}{2}k + \frac{1}{2}l_L k - \frac{1}{2} + \frac{1}{2}l_L \leq \hat{\tau}_L^2(l_L)$ if and only if $l_L \leq 7 - 4\sqrt{3} \approx 0.072$. Hence, the tax constraint binds for all $l_L \geq 7 - 4\sqrt{3}$. Substituting $\hat{\tau}_L^2(l_L)$ into the objective function, we solve the following problem

$$\max_{l_L} \left\{ \left(\hat{\tau}_L^2(l_L) - l_L \right) \left(\frac{1 - l_L k - k + \hat{\tau}_L^2(l_L)}{2 - 2k} \right) \right\}$$

$$\text{s.t. } l_L \in \left[7 - 4\sqrt{3}, \frac{1}{2} \right]$$

Solving this for an interior solution yields two solutions, $l_L^{**1} = 1$ and $l_L^{**2} = \frac{1}{9}$. Only the second is a maximizer. Indeed, we have that $l_L^{**2} = \frac{1}{9} \geq 7 - 4\sqrt{3}$. For simpler notation, we write $l_L^{**2} = l_L^{**}$. Corresponding profits are given by $r_L \left(\hat{\tau}_L^2, l_L^{**}, \tau_F^{s2}, l_F^{s2} \right) = \frac{4}{27}(k - 1)$. One can also verify that a strategy with $l_L < 7 - 4\sqrt{3}$ and no binding tax constraint does not yield more revenue. Again, the derivation is omitted.

It is immediate to see that L prefers the strategy with high standard-provision to the one with low standard provision. It only remains to verify that at this strategy choice of L, Government F indeed wants to set less standard level and set the highest firm sharing tax. Since the tax L sets is, by derivation, the highest one at which F prefers sharing and less provision to undercutting, we only need to verify that F does not want to set more standard and share. But we showed already that this cannot be the case since $l_L^* \geq \frac{1}{2}$. The optimal strategy for L is therefore $(l_L^*, \tau_L^*) = \left(\frac{8}{9}, \frac{4}{3} + \frac{4}{9}k \right)$.

The derivations showing that the strategy choices, if k varies, behave in the way as stated in the proposition are omitted (all omitted parts of the proofs are available upon request). $\square$

Theorem 1. *The outcome of the subgame perfect equilibrium.*

The subgame perfect equilibrium is as follows

a. *(Efficient outcome) If $k \leq \frac{1}{3}$, then $(l_L^*, \tau_L^*) = (0, 0)$, and $(l_F^*, \tau_F^*) = (0, 0)$, and $\hat{s}^* = \frac{1}{2}$.*

b. *(ES haven) If $\frac{1}{3} < k \leq 1$, then $l_L^* \geq \frac{8}{9}$, $\tau_L^* \in (l_L^*, 2l_L^*)$, and $l_F^* = 0$, $\tau_F^* = \frac{2}{3} \left(\sqrt{7k^2 - 6k + 3} - k + 3 \right)(k +$ $\left(\frac{3}{4}, \frac{4}{3} \right)$, and $\hat{s}^* = \frac{1}{3} \left(\sqrt{7k^2 - 6k + 3} - k + 3 \right)(k + 1)^{-1} > \frac{2}{3}$.*

c. *(Excessive ESs) If $k > 1$, then $l_L^* = \frac{8}{9}$, $\tau_L^* = \frac{4}{3} + \frac{4}{9}k > 2l_L^*$, and $l_F^* = \frac{2}{3}$, $\tau_F^* = \frac{4}{3} + \frac{2}{3}k >$ $3l_F^*$, and $\hat{s}^* = \frac{2}{3}$.*

Proof. The subgame perfect equilibrium strategy for Government L is the one derived in Proposition 2. For Government F the outcome is obtained by plugging (l_L^*, τ_L^*) into f^*

as specified in Proposition 1. It is straightforward to verify that, for part b, the taxes lie indeed in the specified range. The equilibrium marginal type of firm is obtained by plugging the equilibrium strategies into $\hat{s}(l_L, \tau_L, l_F, \tau_F)$. Plugging all values into the rent functions yields the corresponding rents. For parts b and c, simple comparison shows that the follower makes higher rents. It is straightforward to verify that $\hat{s}^* > \frac{2}{3}$ in part b. $\square$

References

[1] Andanova, L.B. (2003); *Transnational Politics of the Environment: The European Union and Environmental Policy in Central and Eastern Europe.* MIT Press, Cambridge, MA.

[2] Antweiler, A., B. Copeland and M.S. Taylor (2001); "Is Free Trade Good for the Environment?" *American Economic Review*, 91(4): 877-908.

[3] Baldwin, R. and P. Krugman (2004); "Agglomeration, Integration and Tax Harmonization." *European Economic Review*, 48: 1-23.

[4] Black, D.A. and W.H. Hoyt (1989); "Bidding for Firms." *American Economic Review*, 79(5): 1249-1256.

[5] Boadway, R., K. Cuff and N. Marceau (2004); "Agglomeration Effects and the Competition for Firms." *International Tax and Public Finance* 11: 623-646.

[6] Brennan, G. and J.M. Buchanan, (1980); *The Power To Tax: Analytical Foundations of a Fiscal Constitution,* Cambridge University Press, Cambridge.

[7] Brueckner, J.K. (2000); "A Tiebout/Tax Competition Model." *Journal of Public Economics* 77: 285-306.

[8] Cumberland, J.H. (1979); "Interregional Pollution Spillovers and Consistency of Environmental Policy." Published in H. Siebert et al. eds., *Regional Environmental Policy: The Economic Issues*, New York University Press, New York.

[9] Cumberland, J.H. (1981); "Efficiency and Equity in Interregional Environmental Management." *Review of Regional Studies* 2: 1-9.

[10] d'Aspremont, C., J. Gabszewicz, and J. Thisse (1979); "On Hotelling's Stability in Competition." *Econometrica*, 17: 1145-1151.

[11] Davies, R.B. and C.J. Ellis (2007); Competition in Taxes and Performance Requirements for Foreign Direct Investment." *European Economic Review*, 51(5): 1303-1325.

[12] Ederington, J., A. Levinson and J. Minier (2004); "Trade Liberalization and Pollution Havens." Published in *Advances in Economic Analysis and Policy*, Vol. 4, No. 2, Article 6, the Berkley Electronic Press.

[13] Fischel, W.A. (1975); "Fiscal and Environmental Considerations in the Location of Firms in Suburban Communities." Published in E. Mills and W. Oates (eds.) *Fiscal Zoning and Land Use Controls*, D.C. Heath, Lexington, MA.

[14] Haufler, A. and I. Wooton (1999); "Country Size and Tax Competition for Foreign Direct Investment." *Journal of Public Economics*, 71: 121-139.

[15] Hindriks, J. (1999); "Tax vs. Transfer Competition." *Journal of Public Economics*, 74: 215-234.

[16] Justman, M., J.-F. Thisse, and T. van Ypersele (2002); "Taking the Bite out of Fiscal Competition." *Journal of Urban Economics* 52, 294–315.

[17] Keen, M. and M. Marchand (1997); "Fiscal Competition and the Pattern of Public Spending." *Journal of Public Economics* 66: 33-53.

[18] Kind, H., K. Knarvik and G. Schjelderup (2000); "Competing for Capital in a 'Lumpy' World." *Journal of Public Economics*, 78: 253-274.

[19] Levinson, A. (1997); "A Note on Environmental Federalism: Interpreting Some Contradictory Results." *Journal of Environmental Economics and Management* 33: 359-366.

[20] Levinson, A. (1999a); "NIMBY Taxes Matter: the Case of State Hazardous Waste Disposal Taxes." *Journal of Public Economics* 74: 31-51.

[21] Levinson, A. (1999b); "State Taxes and Interstate Hazardous Waste Shipments." *American Economic Review*, 89(3): 666-677.

[22] Levinson, A. (2003); "Environmental Regulatory Competition: A Status Report and Some New Evidence." *National Tax Journal*, 56(1): 91-106.

[23] Markusen, J.R., E.R. Morey and N. Olewiler (1995); "Competition in Regional Environmental Policies When Plant Locations are Endogenous." *Journal of Public Economics*, 56: 55-77.

[24] Oates, W.E. (1972); *Fiscal Federalism*, Harcourt Brace Jovanovich, New York.

[25] Oates, W.E. and R.M. Schwab (1988); "Economic Competition Among Jurisdictions: Efficiency Enhancing or Distortion Inducing?" *Journal of Public Economics*, 35: 333-354

[26] Post, D. (2002); "Closing the Deception Gap: Accession to the European Union and Environmental Standards in East Central Europe." Published in University of California International and Area Studies (UCIAS) Edited Volume 1: *Dynamics of Regulatory Change: How Globalization Affects National Regulatory Policies*. Berkeley Electronic Press.

[27] Taylor, M.S. (2004); "Unbundling the Pollution Haven Hypothesis." Published in *Advances in Economic Analysis and Policy*, the Berkley Electronic Press.

[28] White, M.J. (1975); "Firm Location in a Zoned Metropolitan Area." Published in E.S. Mills and W.E. Oates (eds.) *Fiscal Zoning and Land Use Controls*, D.C. Heath, Lexington, MA.

[29] Wilson, J.D. (1986); "A Theory of Interregional Tax Competition *Journal of Urban Economics*, 19: 296-315.

[30] Wilson, J.D. (1996); "Capital Mobility and Environmental Standards: Is There a Theoretical Case for a Race to the Bottom?" Chapter 10 in *Fair Trade and Harmonization* by J.N. Bhagwati and R.E. Hudec. MIT Press, Cambridge.

[31] Zissimos, B. and M. Wooders (2008); "Public Good Differentiation and the Intensity of Tax Competition." *Journal of Public Economics*, 92: 1105-1121.

[32] Zodrow, G.R. and P. Mieszkowski (1986); "Pigou, Tiebout, Property Taxation and the Under-Provision of Local Public Goods." *Journal of Urban Economics*, 19: 356-370.