

COMPETITION OVER STANDARDS AND TAXES

by

Valesca Groenert, Myrna Wooders, and Ben Zissimos



Working Paper No. 08-W20

October 2008

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Competition over Standards and Taxes¹

Valeska Groenert²
Vanderbilt University

Myrna Wooders³
Vanderbilt University

Ben Zissimos⁴
Vanderbilt University

This version: October 2008

ABSTRACT: We show that, in competition between a developed country and a developing country over standards and taxes, the developing country may have a ‘second mover advantage.’ A key feature of standards is that, unlike public goods as usually defined, all firms do not unanimously prefer higher 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 ‘standard haven,’ where some firms escape excessively high standards in the developed country; (3) there may be a ‘race to the top’ with standards set excessively high.

KEYWORDS: Fiscal competition, second mover advantage, standards, tax competition.

JEL CLASSIFICATION NUMBERS: H1; H25; H73; H87

¹We would especially like to thank Amrita Dhillon for useful conversations about this paper. We are also grateful for comments from participants at the Workshop on Stability in Competition, CORE, Louvain la Neuve, a European Science Foundation Workshop, Paris 1, the Social Choice and Welfare Meeting in Istanbul and seminars at the University of Warwick and Vanderbilt University. Funding from the Center for the Americas at Vanderbilt University is gratefully acknowledged.

²valeska.n.groenert@vanderbilt.edu

³myrna.wooders@vanderbilt.edu

⁴ben.zissimos@vanderbilt.edu

Dept. of Economics, Vanderbilt University, Nashville, TN 37235.

Tel: ++1 615 322 3339.

Fax: ++1 615 343 8495.

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.⁵ To investigate this issue, our paper develops a model of international competition over standards and taxes. By a ‘standard’ we have in mind such things as labor regulations, pollution control and property rights enforcement. Firms who locate in a country are required to pay taxes which are used, at least in part, to enforce the standard in that country. The main purpose of this paper is to show that, through competition in standards 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.⁶

Although often modeled as a type of local public good, standards have an important distinguishing feature. A reasonable assumption in the context of most public goods is that (for a given tax outlay) all firms at least weakly prefer a higher level of public good provision. On the other hand all firms do not unanimously prefer higher standard levels. For example, a high level of property rights enforcement may benefit a firm engaged in research and production of new pharmaceutical products while it may hurt a firm engaged in the mass-production of generic drugs. To an otherwise familiar model of fiscal competition, we introduce the assumption that firms have differing ideal standard levels. It is this assumption that gives rise to the second-mover advantage in standard setting that we identify.

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 (with Niskanen 1977 as its source) national governments

⁵For example, although EU accession requirements demand moves towards harmonization of environmental standards and some measures have been put 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. Andonova (2003) provides extensive details of these environmental standards. Although environmental standards provide a good motivating example, our concern will be with standards more broadly defined as we shall explain.

⁶This issue has been raised both with respect to developing countries and to countries from the former Soviet Union often referred to as ‘transition countries/economies.’ For brevity, throughout paper we will use ‘developing country’ as a catch-all term.

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). 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' 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 which depend on the cost of standard mismatch. There are three possible outcomes. (1) If the cost of standard mismatch is low then tax competition leads to an efficient equilibrium outcome (as in Brennan and Buchanan's 1980 model of tax competition). (2) If the cost of standard mismatch is in an intermediate range then the developed country sets standards inefficiently high and the developing country becomes a *standard haven*; a place where firms that prefer a low standard locate in order to escape the high standard set in the developed country. It is especially interesting that inefficiently high standards in the developed country arise in equilibrium purely through strategic interaction between governments in their competition for firms and not as a result of attempts by governments to win the favor of a voting public. (3) If the cost of standard mismatch is high then there is a *race to the top*; both governments set standards inefficiently high and, because countries are differentiated by their standard levels, the intensity of tax competition is reduced as well. The precise set of interactions that gives rise to these equilibrium outcomes will be described in due course.

Much of the literature that examines fiscal competition where the public good in question is a standard assumes that (for a given tax take) citizens at least weakly prefer higher standards and that the standard in question is environmental. As a result, insights from the literature on tax competition with local public goods extend naturally; see Wilson (1996) for a survey. Broadly, the literature can be categorized into three areas. The first category, following Tiebout (1956), focuses on situations where competition among independent governments is like competition among firms and enhances efficiency. The second category concerns the presence of a policy-failure that allows or induces governments to set taxes on capital too high. This in turn induces governments to try to offset the depressive effects of capital taxes on investment by setting environmental standards too lax; this outcome is

popularly known as a ‘race to the bottom.’ See Oates and Schwab (1988) for further details, as well as a discussion by Wilson (1996) of Oates and Schwab plus the related literature. The third category considers situations in which there is strategic interaction, over standards and taxes, between governments and a small number of firms. See for examples Markusen, Morey and Olewiler (1995) and Davies and Ellis (2007). In such settings, strategic interactions over the market power held by firms and the policy failures of governments are the source of inefficient policy choices.

Our model combines features of models from papers in the first two categories: 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 market - or policy - failures that preclude the attainment of a fully efficient equilibrium. 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 contrast to our model, Fischel (1975) and White (1975) 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. 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. Moreover, there is policy failure in the sense that once the levels of public goods – in our model, standards – are fixed they cannot be altered. Another difference is that we fix the number of countries (at two) which enables their rent-maximizing governments to make positive rents and thus allows inefficiencies to arise.⁷

Rent-maximizing governments are a source of policy inefficiency for Oates and Schwab (1988) as well. Again, if governments are able to earn rents from taxation of mobile resources (in their case, capital) then there is an incentive to simultaneously set standards inefficiently low. Other papers in the literature build on these basic features in various ways. Interestingly, although the source of excessive taxation put forward by Oates and Schwab is the same as ours, their outcome in terms of environmental standards is starkly different. In their setting the outcome is a race to the bottom; in our setting, if the marginal cost of

⁷Our focus is on national governments while in much of the literature on standards and tax competition governments preside over jurisdictions more broadly defined. The reason we interpret the context of our model as international is that the range of policy options under consideration is more limited than in a domestic or federal context. In particular, the feature of our model that transfers between jurisdictions are not allowed appears to mirror more closely an international setting.

standard mismatch is sufficiently large, the outcome is a race to the top.⁸

The remainder of this paper is organized as follows. Section 2 introduces the model, then defines strategies and the subgame perfect equilibrium. Section 3 solves for the efficient allocation. Section 4 presents the equilibrium outcome, which is defined in terms of the location decisions of firms and policies set by the developing country and the developed country respectively. Conclusions are drawn in Section 5.

2 The Model

The governments of two countries, a developed country, L (for ‘leader’), and a developing country, F (for ‘follower’), compete over standard 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 standard in the country where it locates. We will first specify the behavior of firms, and then we will turn to the 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 .⁹ The tax levied on the firm is τ_L if it locates in L and τ_F if it locates in F . The value $s \in [0, 1]$ uniquely identifies a firm and its ideal standard level.¹⁰ The *standard mismatch* for a firm s is given by the difference between s and the standard level actually set in the country where the firm locates. The impact of standard mismatch on production costs is parameterized by k ; we refer to k as the *marginal cost of standard mismatch*. If we let the variables $l_L, l_F \in [0, 1]$ denote the standard levels set by L and F respectively then we can express the profit function for firm s as follows:

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

⁸Wilson (1996) insightfully conjectures that, under certain parameterizations, it may be possible to show that Oates and Schwab’s framework could motivate a race to the top as well.

⁹To 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.

¹⁰We choose the interval $[0, 1]$ to simplify the exposition. The same qualitative results may be obtained using an arbitrary interval $[a, b]$.

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 will serve as an upper bound for the tax that a government can set.

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

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

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:

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

Firm s may prefer one country, say F , in terms of the tax that it sets; $\tau_F < \tau_L$. But if L 's standard is sufficiently close to s (i.e. $|s - l_L| < |s - l_F|$) then L can attract s to its country.¹¹ If there is more than one firm that is indifferent between the two countries, 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 $\tau_L + kl_L = \tau_F + kl_F$. 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 countries defined in Section 2.2.¹²

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

- (1) Given taxes and standards, firms that are indifferent between the two countries locate in the country with the lower standard mismatch, i.e. firms care more about the persistence of an established standard level than the constancy of a given tax level;
- (2) If all firms are indifferent between the two countries, then half locate in one country and half locate in the other;
- (3) If a government has multiple best responses, it chooses the best response that maximizes its share of firms.

These assumptions will be discussed further when we derive equilibrium in Section 4. The location decisions of firms described above are illustrated in Figure 1.

¹¹Firms' 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.

¹²See the Appendix for additional details.

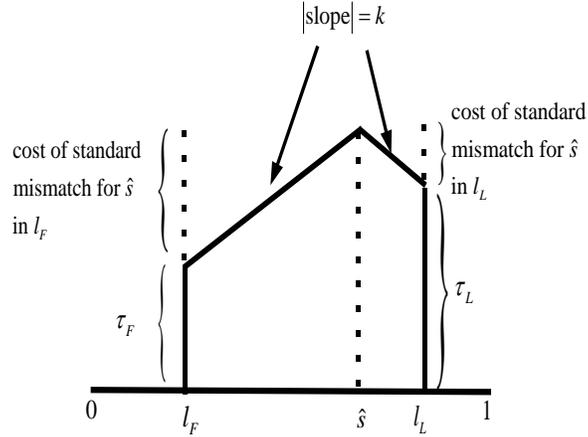


Figure 1

Figure 1 is reminiscent of ‘Hotelling’s umbrella,’ and reflects the underlying structure of our model which has some Hotelling features (see Hotelling 1929). The figure shows illustrative levels of standards and taxes set by governments F and L . For standards and taxes as shown, the point \hat{s} represents the ideal standard 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 standard provision. A government’s cost of enforcing a standard level $l \in [0, 1]$ is l per firm that is located in its country. Thus the cost of enforcing a standard is assumed to be proportional to the level of the standard and the number of firms over which it must be enforced. Government F takes l_L and τ_L as parameters and chooses l_F and τ_F to maximize its rents. Discontinuities arise in the 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:

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

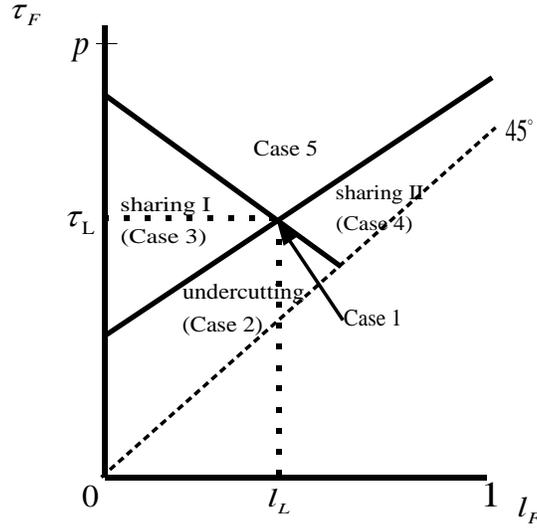


Figure 2

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 standard 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 standard 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 standard than L in Case 3 and a higher standard 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 made by firms plus the governments' rents. An allocation consists of two standard levels and an assignment of firms to countries, denoted by (l_F, l_L, \hat{s}) . Formally, the allocation (l_F, l_L, \hat{s}) is efficient if it solves

$$\begin{aligned} & \max_{\{l_F, l_L, \hat{s}\}} \int_0^{\hat{s}} (p - c - \tau_F - k |l_F - s|) ds + (\tau_F - l_F) \hat{s} \\ & + \int_{\hat{s}}^1 (p - c - \tau_L - k |l_L - s|) ds + (\tau_L - l_L) (1 - \hat{s}) \\ & \text{s.t. } l_F \in [0, 1], l_L \in [l_F, 1], \text{ and } \hat{s} \in [0, 1]. \end{aligned}$$

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

$$\begin{aligned} & \min_{\{l_F, l_L, \hat{s}\}} \int_0^{\hat{s}} k |l_F - s| ds + l_F \hat{s} + \int_{\hat{s}}^1 k |l_L - s| ds + l_L (1 - \hat{s}) \\ & \text{s.t. } l_F \in [0, 1], l_L \in [l_F, 1], \text{ and } \hat{s} \in [0, 1]. \end{aligned}$$

Thus the efficient allocation minimizes the sum of the aggregate costs of standard mismatch and the costs of standard provision. 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 standard 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 standard level l that would lower mismatch costs: $k < l + k(1 - l) = k + l(1 - k)$.

If $k > 1$, the results are not that obvious. Solving the maximization problem yields the efficient allocation:

$$\begin{aligned} l_F^e(k) &= \frac{(3k - 2)(k - 1)}{(13k - 7)2k}; \\ l_L^e(k) &= \frac{(3k - 2)(k - 1)}{(13k - 7)2k} + \frac{(3k - 2)}{(13k - 7)}; \\ \hat{s}^e(k) &= \frac{(3k - 2)}{(13k - 7)}. \end{aligned}$$

All values are increasing in k . Figure 3 illustrates the efficient standard levels and the allocation of firms to countries depending on k for the case $k > 1$. As $k \rightarrow \infty$, we have $l_F^e(k) \rightarrow 0.12$, $l_L^e(k) \rightarrow 0.35$, and $\hat{s}^e(k) \rightarrow 0.23$.

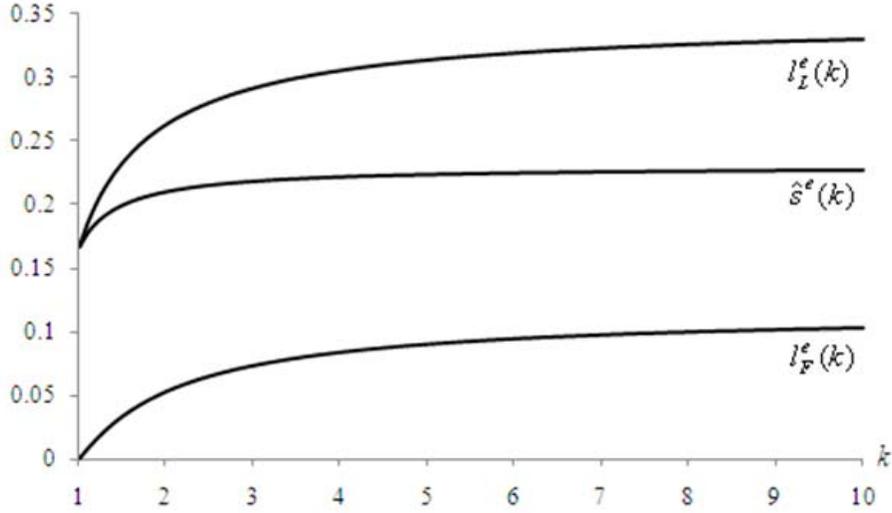


Figure 3

For $k > 1$, in the efficient allocation most firms are assigned to the country with the higher standard provision. However, standard provision is relatively low in both countries, at most about a third of the maximum possible standard $s = 1$. A comparison with our equilibrium results will show that, except for the case when $k \leq \frac{1}{3}$, there is overprovision of the standard in equilibrium. 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). To understand our efficient solution, note that if our model were a direct application of Hotelling's model the level of standard provision would not have affected its cost of provision. In our model, when determining the efficient outcome one has to take into account the costs of providing the standard for each firm assigned to a country as well as the costs of standard mismatch; in Hotelling's model only the costs of mismatch matter. Because higher standard levels are more expensive, the efficient standard levels are lower than they would have been in a direct application of Hotelling's model.

4 Competition over Standards and Taxes

In this section we will derive and discuss the equilibrium outcome. Our approach will be to first define equilibrium and then state our main theorem in which equilibrium is characterized. The derivation of equilibrium will be undertaken subsequently.

As mentioned above, standard provision and tax setting are modeled as a two-stage game. The sequence of events is as follows. Government L sets its standard level and tax and then, observing L 's choices, Government F sets its standard 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 a standard level and a tax. A strategy is *feasible* if the tax is high enough to cover the cost of standard provision.¹³ Formally, the set of feasible strategies is

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

A strategy for Government F is a mapping that assigns a pair, consisting of a standard 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 $f : S_L \rightarrow [0, 1] \times [0, p]$ where $f(l_L, \tau_L) = (l_F, \tau_F)$. Let F be the set that contains all such mappings. The set of feasible strategies for Government F consists of those members of F with the property that tax revenue covers the cost of the associated standard level; that is,

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

We are interested in the pure strategy subgame-perfect Nash equilibrium of the game, which can be viewed as a Stackelberg game.¹⁴

Definition 2. *A pure strategy subgame-perfect Nash equilibrium in taxes and standard 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.

¹³Thus we make the simplifying assumption that there are no other sources of government revenue and no international capital market which governments can tap. We do not think that allowing such a possibility would change our results, wherein governments make positive rents in equilibrium.

¹⁴It 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.*¹⁵

The subgame-perfect equilibrium is as follows.

- a. *(Efficient outcome) If $k \leq \frac{1}{3}$, both countries set the minimum standard 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. *(Standard haven) If $\frac{1}{3} < k \leq 1$, the differentiation in standard levels between the two countries is high; the developed country sets a standard close to the maximum level and the developing country sets a zero standard level. Both countries set taxes that lead to positive rents, and rents are always higher for the developing country than for the developed country. The majority of firms locates in the developing country. Specifically, it holds that $l_L^* \geq \frac{8}{9}$, $\tau_L^* \in (l_L^*, 2l_L^*)$, and $l_F^* = 0$, $\tau_F^* \in (\frac{3}{4}, \frac{4}{3})$, and $\hat{s}^* > \frac{2}{3}$.*
- c. *(Race to the top) If $k > 1$, the standard level is above $\frac{1}{2}$ in both countries, with the developed country setting a higher standard than the developing country. The standard levels do not vary with k . Both governments make positive rents, requiring firms to pay more than twice the cost of standard provision. The developing country sets a higher tax than the developed country and earns higher rents. Two-thirds of the firms locate in the developing country, and every firm with strictly higher ideal standard level than set in the developing country locates in the developed country. 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 standard levels set in the two countries depending on k . The subgame-perfect standard and tax levels differ considerably across the three regions of k : A small k leads to an efficient outcome; for k in an intermediate range there is almost maximum differentiation in standards; for large k there is some differentiation but it is substantially smaller than for k in the intermediate range. The reason is that F sets two-thirds of the maximum standard instead of zero as in the ranges where k is low and high. For low k , taxes are the same in both countries. For $\frac{1}{3} < k \leq 1$, the developed country sets a higher tax than the developing country, whereas for $k > 1$, the developing country sets a higher tax than the developed country.

¹⁵The theorem is restated in the Appendix with formulae for all the equilibrium values shown explicitly.

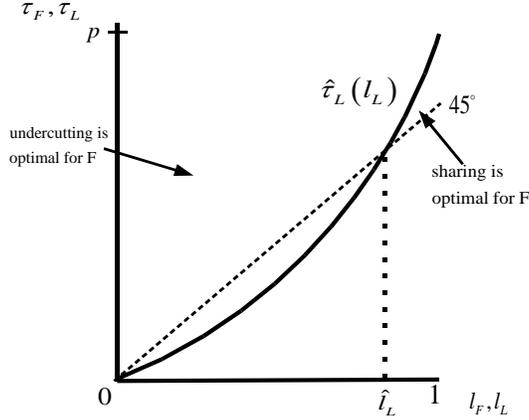


Figure 4

The common characteristic of equilibrium across all levels of k is that the developing country attracts at least as many firms as the developed country. Also, for $k > \frac{1}{3}$, both the developed and developing country are able to extract rents.¹⁶ 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. An additional interesting aspect is that F , who sets a lower standard, makes more rents because it both attracts more firms and makes more rents per firm. Except for $k \leq \frac{1}{3}$, countries set inefficiently high standard levels.

The intuition behind the result for low marginal cost of standard mismatch, case (a), is straightforward. For $k \leq \frac{1}{3}$, the costliness of standard mismatch is so low that countries do not succeed in differentiating themselves via standard levels. This is due to the fact that firms do not perceive countries with different standard levels as sufficiently distinct from each other. Therefore a country cannot extract a monopolistic rent by setting a standard level different from the one set in another country. All competition occurs in taxes, which brings about an efficient outcome.

Turning to case (b), the intuition behind the maximum differentiation in standards that occurs when the marginal cost of standard mismatch is in an intermediate range is as follows. The developing country has a second-mover advantage and so creates a standard haven for firms whose costs are affected more by taxes than by standard mismatch. The developed

¹⁶The result is particularly striking for the country that supplies zero standard even though it levies a positive tax.

country can extract some rents (because k is not too small), but only by differentiating itself substantially (because k is not too large) from the developing country. Because it is a dominant strategy for the developing country to become a standard haven, the developed country can only differentiate itself by setting its standard at a high level. As a result there is close to maximum differentiation between the two countries.

Regarding case (c), when the cost of standard mismatch is high relative to taxes, both countries offer inefficiently high standard levels. (Recall from Section 3 that the efficient outcome calls for the countries offering up to, respectively, 12% and 35% of the maximum standard level.) Because firms value a lower standard mismatch more than lower taxes, the developing country has an incentive to choose a standard level close to the standard level that the developed country sets. Whether the developing country chooses a standard level that is lower or higher than the one in the developed country depends on whether the developed country sets a relatively high standard level (in which case the developing country would set a lower standard level) or whether it sets a relatively low standard level (in which case the developing country would set a higher standard level). In equilibrium, the developed country chooses a high standard level even though a lower standard level would be less costly. This is because the developed country has to allow the developing country to extract high rents to prevent the developing country from undercutting.

Notice that case (c) is the case which one would expect to be least stable among the three cases. Because F sets the highest tax that still attracts a positive fraction of firms to its country (all firms ‘to the left of F ’ with an ideal standard level not higher than the one F sets), Government L - if able to do so - could marginally lower its tax, and by doing so attract all firms to its country. An additional fraction of two-thirds of all firms would be attracted, from which L could extract rents.

Now that we have stated our main result and given the basic intuition behind it, we will next provide a detailed analysis of its derivation. To do so, the next subsection provides a characterization of L ’s best response, and this is followed by a characterization of F ’s best response in the subsection that follows. All proofs are given in the Appendix.

4.1 The developing country’s best response function

In this section, we analyze Government F ’s best response to a given strategy (l_L, τ_L) of Government L . 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 .

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*.

Government F 's optimization problem.

(a) *Maximize rents over sharing I*

$$\begin{aligned} & \max_{\tau_F, l_F} (\tau_F - l_F) \hat{s} \\ & \text{s.t.} \\ & l_F \in [0, l_L) \\ & \tau_F \in [l_F, p] \\ & \tau_F \in [\tau_L - k(l_L - l_F), \tau_L + k(l_L - l_F)] \end{aligned}$$

(b) *Maximize rents over sharing II*

$$\begin{aligned} & \max_{\tau_F, l_F} (\tau_F - l_F)(1 - \hat{s}) \\ & \text{s.t.} \\ & l_F \in (l_L, 1] \\ & \tau_F \in [l_F, p] \\ & \tau_F \in [\tau_L - k(l_F - l_L), \tau_L + k(l_F - l_L)] \end{aligned}$$

(c) *Maximize rents over undercutting*

$$\begin{aligned} & \max_{\tau_F, l_F} (\tau_F - l_F) \\ & \text{s.t.} \\ & l_F \in [0, 1] \\ & \tau_F < \tau_L - k|l_L - l_F|. \end{aligned}$$

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.

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.¹⁷

In our model this difficulty can be resolved in a straightforward way. Even though one cannot determine an optimal undercutting strategy, one can determine when Government F will undercut L and when it will share firms with L . Because Government L will avoid strategies that induce F to undercut (i.e., undercutting happens only off the equilibrium path), we can solve our model without determining the specific undercutting strategy. We determine which of L 's strategies lead F to undercut by assuming that F undercuts whenever there exists *some* undercutting strategy that yields more rents than the best sharing strategy.

To be more specific, let $r_F^s(l_L, \tau_L)$ be F 's rent from an optimal sharing strategy after L has chosen (l_L, τ_L) , 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 $r_F^u(l_L, \tau_L) = \lim_{\varepsilon \rightarrow 0} r_F^u(l_L, \tau_L; \varepsilon)$. Note that, by choosing ε sufficiently small, F can obtain a rent arbitrarily close to $r_F^u(l_L, \tau_L)$, but still $r_F^u(l_L, \tau_L; \varepsilon) < r_F^u(l_L, \tau_L)$ no matter how small is ε . By solving $r_F^s(l_L, \tau_L) = r_F^u(l_L, \tau_L)$ we obtain a critical tax τ_L that depends on l_L . We denote this tax by $\hat{\tau}_L(l_L)$ and will refer to it as the *sharing tax limit*. The sharing tax limit can be used to classify payoffs to F 's standard and tax as follows:

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

In other words, if L 's tax is higher than the sharing tax limit, then F can find an ε small enough to make the rents earned from undercutting higher than the rents earned by sharing. However, if L sets its tax no higher than the sharing tax limit, F finds that sharing yields strictly higher rents than undercutting, no matter how small is ε . Figure 5 depicts the situation.

¹⁷The 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.

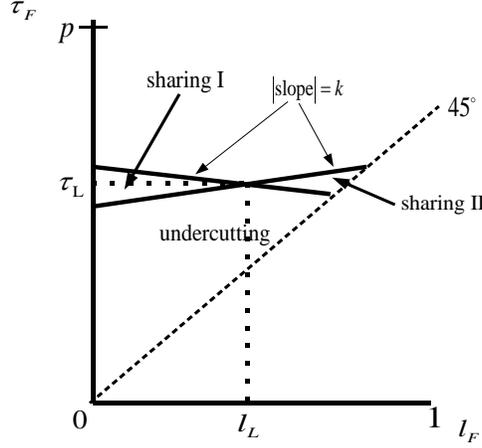


Figure 5

To deal with the fact that Government F might have multiple best responses recall our assumption that, if a government has multiple best responses, it chooses the best response that maximizes its share of firms. This implies that of the best responses available, F chooses the one that requires the lowest level of standard. Moreover, if in addition $(l_L, \tau_L) = (0, 0)$, we assume that F sets $\tau_F = 0$. We only require these properties in two situations. First, if $k < 1$ and $(l_L, \tau_L) = (0, 0)$, there is no response that yields positive rents for F . Our assumption then implies that F chooses $(l_F, \tau_F) = (0, 0)$. Notice that any other feasible strategy for F would induce all firms to locate in L . Second, if $k = 1$ then for any (l_L, τ_L) Government F has a whole range of best responses. More specifically, there is a best response at each standard level l_F . The reason is that standard mismatch and taxes are equally costly for firms. Therefore if a country decides, for example, to set a lower standard and use the resources that it saves to reduce its tax, it will attract the same share of firms as before and it will make the same rents per firm. For this case, our assumption implies that $l_F = 0$.

Before stating our first result, 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 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 $\{(l_F^u, \tau_F^u), (l_F^{s1}, \tau_F^{s1}), (l_F^{s2}, \tau_F^{s2})\}$.

The nature of the results we obtain differs across three intervals, $k \in (0, \frac{1}{3}]$, $k \in (\frac{1}{3}, 1]$, and $k \in (1, \infty)$. For each of the three regions of k , Proposition 1 summarizes the best

response of Government F to any standard level and tax that Government L has chosen in the first stage.

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 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).

Part (a) of Proposition 1 shows that for small k undercutting dominates sharing. All firms can be attracted without having to set the tax much below L 's tax. At the same time, the area in policy space over which firms are shared is reduced - as k is reduced, a given tax set by F will induce all firms to locate in country L . Figure 6 illustrates the situation. A reduction of k increases undercutting possibilities while at the same time it reduces sharing possibilities. In particular, as $k \rightarrow 0$, the set of sharing possibilities shrinks to the empty set. 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 $\tau_F = 0$.

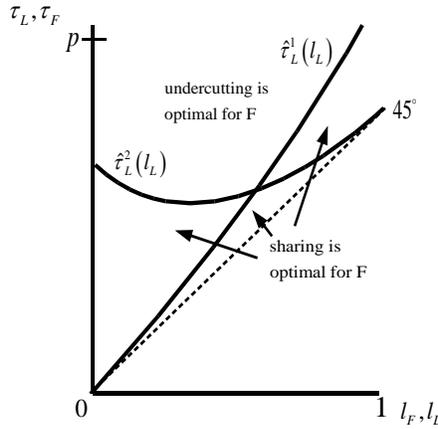


Figure 6

Part (b) of Proposition 1 illustrates the role of the sharing limit tax, $\hat{\tau}_L(l_L)$, in the model. Government F shares if L 's tax does not exceed $\hat{\tau}_L(l_L)$ but undercuts if L 's tax is above it. To see why suppose that, for some standard and tax levels, L and F are sharing firms. If L increases its tax, F will raise its own tax by only half the amount ($\tau_F^{s1} = \frac{1}{2}\tau_L + \frac{k}{2}l_L$). 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, when $k \leq 1$, 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. 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 $s \leq l_L$ will locate in F . At this point, F will want to switch to an undercutting strategy because F has to lower τ_F only marginally to induce an additional share of $(1 - l_L)$ firms to locate in its country.

Similarly, there exists a sharing tax limit in part (c). In Figure 7 the sharing tax limit is given by $\hat{\tau}_L(l_L) = \max \{ \hat{\tau}_L^1(l_L), \hat{\tau}_L^2(l_L) \}$, where $\hat{\tau}_L^1(l_L)$ is the tax up to which an optimal sharing I strategy is better for F than any undercutting strategy and $\hat{\tau}_L^2(l_L)$ is the tax up to which an optimal sharing II strategy is better for F than any undercutting strategy. 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^{s1}, τ_F^{s1}) or (l_F^{s2}, τ_F^{s2}) 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 will yield higher rents for F . More specifically, we show in the appendix that if $l_L \geq \frac{1}{2}$ then setting an even higher standard and sharing is never a best response for Government F . In this case, instead of setting a standard that exceeds L 's standard by x , i.e., $l_F = l_L + x$, and set some tax, F can set $l_F = l_L - x$, without changing the tax. Doing so will increase F 's rent per firm, and the share of firms attracted will be at least as large.

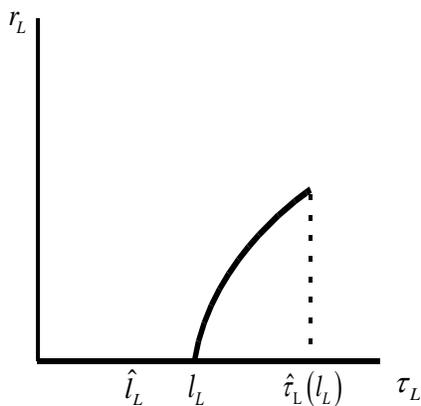


Figure 7

4.2 The developed county'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 . Formally L 's problem is

$$\begin{aligned} & \max_{\{l_L, \tau_L\}} r_L(l_L, \tau_L, f^*(l_L, \tau_L)) \\ & \text{s.t.} \\ & l_L \in [0, 1] \\ & \tau_L \in [l_L, p] \end{aligned}$$

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 would induce F to undercut and hence would leave L with zero rents. Accounting for F 's response in the second stage, those taxes will yield zero rents for L , while a tax that induces F to share firms yields positive rents. Thus, we can formulate a reduced problem for L .

Government L 's reduced optimization problem.

$$\begin{aligned} & \max_{\{l_L, \tau_L\}} r_L(l_L, \tau_L, f^*(l_L, \tau_L)) \\ & \text{s.t.} \\ & l_L \in [0, 1] \\ & \tau_L \in [l_L, \hat{\tau}_L(l_L)] \end{aligned}$$

Figure 8 depicts L 's rent function, $r_L(l_L, \tau_L, f^*(l_L, \tau_L))$, 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 $\tau_L = \hat{\tau}_L(l_L)$.

As with Government F , Government L 's best response to f might not be unique. Government L 's best response to F 's equilibrium strategy f^* is not unique if and only if $k \leq \frac{1}{3}$. In this case L cannot make any positive rents because F always undercuts L . Paralleling our assumption for Government F , we assume that L then chooses $(l_L, \tau_L) = (0, 0)$. As before, this assumption reflects a preference for strategies that attract larger shares of firms.

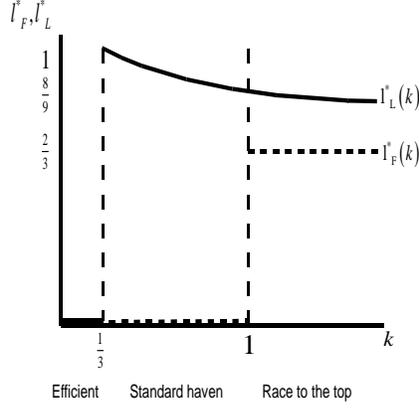


Figure 8

Since the nature of f^* depends on k , so will the nature of Government L 's optimal strategy; Proposition 2 summarizes. 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 5 for an illustration.

Proposition 2.¹⁸ (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, $(l_L^*, \tau_L^*) = (\frac{8}{9}, \frac{4}{3} + \frac{4}{9}k)$.

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 assumption, Government L picks $(l_L^*, \tau_L^*) = (0, 0)$. Rents for both governments are zero.

¹⁸Proposition 2 is restated in the Appendix with the exact expressions for the optimal strategies.

To derive the results for $\frac{1}{3} < k \leq 1$, we show that for $k \leq 1$, 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 $l_L > \hat{l}_L$, as a result of which we can focus on standard levels $l_L \in [\hat{l}_L, 1]$. From this we can see immediately that L chooses the tax $\hat{\tau}_L(l_L)$ since, as shown in Figure 5, L 's rents are increasing in τ_L up to $\hat{\tau}_L(l_L)$.

At first sight it seems surprising that for $\frac{1}{3} < k \leq 1$ only a standard of more than $\hat{l}_L = \frac{8k(1-k)}{(1+k)^2}$ allows L to earn positive rents (see Lemma 4 in the Appendix). On the face of it, there is of course an incentive for L to set a low standard level since this saves monitoring costs and would increase the share of firms locating in L . However, the lower the standard level that L sets, the greater the incentive for F to switch to an undercutting strategy because switching from a sharing strategy to an undercutting strategy 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. This happens despite the fact that standard mismatch is not very important to firms.

When $k > 1$, L sets a high standard level. Intuitively, for L , setting a low standard level in order to induce F to set a higher standard level might seem a better strategy. However, it is in fact better for L to let F be the country that sets a low standard. This guarantees F higher rents from sharing, which 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 $\tau_L = \hat{\tau}_L(\frac{1}{9})$ instead of its actual equilibrium choice $l_L^* = \frac{8}{9}$ and $\tau_L^* = \hat{\tau}_L(\frac{8}{9})$. 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, $\hat{\tau}_L(\frac{1}{9})$, is so much lower than $\hat{\tau}_L(\frac{8}{9})$ 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. 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.

Notice also that, for $k > 1$, in contrast to the situation where $k \leq 1$, the standard level set by L does not vary with k . As standard mismatch becomes more costly for firms, L extracts more rents through an increase in taxes.

It should now be clear that 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 above.

5 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 advantage in setting standards and taxes. This paper sets out a formal framework which makes precise a set of circumstances under which such a second mover advantage may arise. Three possible predictions are made about the outcome of fiscal competition when the public good in question is a standard. The particular prediction that emerges in equilibrium depends on the marginal cost of standard mismatch. The model focuses on the interplay between governments' incentives to manipulate policy - standards and taxes - in order to maximize rents and firms' incentives to locate where these policies have the most favorable impact on profits. The key point is that the government of the developed country wants to avoid inducing the developing country to undercut because being undercut implies losing all firms and hence all rents. If the marginal cost of standard mismatch is low, then standards 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 are strong enough to 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 standard relative to the tax. A proportion of firms will 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 standards too high.

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. While our work addresses the issue of 'standard havens' more broadly defined, it is in the area of the environment that the idea of a haven has attracted the most attention and so it seems worth evaluating the contribution of our work in that context. 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 environmental standards. 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 environmental standards are a source of comparative advantage since they make the

opportunity cost of pollution 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 solely on the feature that developed countries have tended to introduce environmental standards earlier than developing countries.

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 standards ‘too low.’ While it seems reasonable to argue that developed countries may set standards too high, a concern is that developing countries actually set their standards too low from the perspective of consumers. The introduction of consumers to the model could make it possible for standards 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 n th 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.¹⁹

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

¹⁹It is tempting to think that one could analyze a model in which a ‘core’ country sets policy first and a larger number of periphery countries set policy subsequently (but at the same time as each other). However, the difficulty here is that in the present framework in general there may not exist an equilibrium in pure strategies when countries set policies simultaneously.

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 standard level and they would set taxes at the level of prices, 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 standards; for $k > 1$ the efficient outcome does have a positive level of standard 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 may be overwhelming. This may explain why in practice proposals for collusion have tended to be weaker, entailing for example the introduction of minimum standards. A surprising implication of our framework is that it is not in the interest of the developed country to introduce a binding minimum standard. The reason is that the developed country benefits from being able to differentiate itself from the developing country and putting in place a minimum standard would limit the scope for doing so. Thus our model presents a possible way of understanding situations in which standards have been called for but none have actually emerged.

A Appendix

A.1 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.2 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 $\tau_F = \tau_L - kl_L - \varepsilon \geq 0$ 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 $(l_F^u, \tau_u^u) = (0, \tau_L - kl_L - \varepsilon)$ 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 $\tau_F' = \tau_L - kl_L - \varepsilon$ is feasible (i.e., $\tau_L - kl_L - \varepsilon > 0$) and

yields more rents per firm because it saves costs of l_F per firm and reduces revenue per firm by at most kl'_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 $(l_F^u, \tau_F^u) = (l_L, \tau_L - \varepsilon)$ 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

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

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

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

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 $(l'_F, \tau'_F) = (0, \tau_F - kl_F)$. 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 $(l_F, \tau_F) \in \text{sharing I}$. 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 $(l_F, \tau_F) \in \text{sharing II}$ is analogous. \square

Lemma 2b.

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

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

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

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

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,

$$\begin{aligned} & \max_{\tau_F} \left\{ \tau_F \left(\frac{\tau_L - \tau_F}{2k} + \frac{l_L}{2} \right) \right\} & ((*) \\ \text{s.t. } & \hat{s}(l_L, \tau_L, 0, \tau_F) \in [0, l_L] \\ & \tau_F \geq 0 \end{aligned}$$

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

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

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 of 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 $\tau_L > kl_L$.*

Proof: For $k \leq \frac{1}{3}$, we have $\tau_L > 3kl_L$ for all strategies with $\tau_L > kl_L$. 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.²⁰ 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 than 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

$$\begin{aligned} \left(\frac{1}{2}\tau_L + \frac{k}{2}l_L\right) \left(\frac{1}{4k}\tau_L + \frac{l_L}{4}\right) &= (\tau_L - kl_L) \iff \\ kl_L - \tau_L + \frac{1}{8k} (2kl_L\tau_L + \tau_L^2 + k^2l_L^2) &= 0 \end{aligned}$$

²⁰Notice 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

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

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

$$\begin{aligned} & \max_{(l_F, \tau_F)} \left\{ (\tau_F - l_F) \left(\frac{\tau_L - \tau_F}{2k} + \frac{l_L + l_F}{2} \right) \right\} \\ & \text{s.t.} \\ & \tau_F \geq l_F \\ & l_F \in [0, l_L] \\ & \tau_F \in [\tau_L - k(l_L - l_F), \tau_L + k(l_L - l_F)] \end{aligned}$$

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

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

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

for $\delta > 0$.

Therefore, we can reduce F 's problem to

$$\begin{aligned} \max_{l_F} \{ &(\tau_L + k(l_L - l_F) - l_F) l_F \} \\ \text{s.t.} & \\ l_F \in & [0, l_L) \end{aligned}$$

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

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

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

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

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 would set would be $\tau_F = \tau_L = l_L(k+2)$. By assumption, it would attract half of the firms and therefore

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

for small enough 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 notice that

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

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 undercutting²¹:

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

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.

²¹As 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

$$\begin{aligned} r_F^{s2}(l_L, \tau_L) - r_F^u(\tau_L, l_L) &= \frac{1}{4} (k-1)^{-1} (k - l_L k - 1 + \tau_L)^2 - (\tau_L - l_L) = 0 \\ \iff \hat{\tau}_L^2(l_L) &= -z + l_L k + k - 2(k-1) \sqrt{l_L} \end{aligned}$$

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 2.²² *(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 $(l_L^*, \tau_L^*) = \left(\frac{2k}{9} (k+1)^{-2} \frac{1 - \left(\frac{1}{3k+3} (4k - \sqrt{3 - 6k + 7k^2})\right)^2}{(k^2 + 6\sqrt{7k^2 - 6k + 3} + 2k\sqrt{7k^2 - 6k + 3} + 15)} \right)$.

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

²²Proposition 2 and Theorem 1 are restated here in the Appendix with the exact expressions for the optimal strategies.

reduced optimization problem for L is therefore

$$\begin{aligned} & \max_{(\tau_L, l_L)} \{(\tau_L - l_L) (1 - \hat{s}(\tau_L, l_L, \tau_F^s, 0))\} \\ \text{s.t. } & l_L \in [0, 1] \\ & \tau_L \in [l_L, \hat{\tau}_L(l_L)] \end{aligned}$$

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

$$\begin{aligned} & \max_{l_L} \\ & \left\{ \left(k \left(4 - 4\sqrt{1-l_L} - l_L \right) - l_L \right) \left(1 - \frac{\frac{1}{2}(k(4 - 4\sqrt{1-l_L} - l_L)) - \frac{k}{2}l_L}{2k} - \frac{l_L}{2} \right) \right\} \\ \text{s.t. } & l_L \in (0, 1) \end{aligned}$$

The solution to which is $l_L^* = 1 - \left(\frac{1}{3k+3} (4k - \sqrt{3 - 6k + 7k^2}) \right)^2 \in (0, 1)$. 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 $\tau_L > \hat{\tau}_L^2(l_L)$. It follows that $f^*(l_L, \tau_L) = (l_F^u, \tau_F^u)$ if $\tau_L > \max[\hat{\tau}_L^1(l_L), \hat{\tau}_L^2(l_L)]$. 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 $\tau_L > \max[\hat{\tau}_L^1(l_L), \hat{\tau}_L^2(l_L)]$ as those yield zero rents, while the strategy $(l_L, \tau_L) = (0, \hat{\tau}_L^2(0))$ 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 $l_L \in [0, \frac{1}{2}]$, $r_F^{s2} \geq r_F^{s1}$, 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 $l_L \in [0, \frac{1}{2}]$, 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 $\tau'_F = \tau_F$.

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

Under this supposition, government L 's problem is

$$\begin{aligned} & \max_{(\tau_L, l_L)} \{(\tau_L - l_L) (1 - l_F^{s1})\} \\ & \max_{(\tau_L, l_L)} \left\{ (\tau_L - l_L) \left(1 - \frac{kl_L + \tau_L}{2(k+1)} \right) \right\} \\ & \text{s.t. } l_L \in [0, 1] \\ & \tau_L \in [l_L, \hat{\tau}_L^1(l_L)] \end{aligned}$$

Solving for an interior solution yields $\tau_L = 1 + \frac{1}{2}l_L - \frac{1}{2}l_Lk + k$. But $1 + \frac{1}{2}l_L - \frac{1}{2}l_Lk + k \leq \hat{\tau}_L^1(l_L)$ 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

$$\begin{aligned} & \max_{l_L} \left\{ (\hat{\tau}_L^1(l_L) - l_L) \left(1 - \frac{kl_L + \hat{\tau}_L^1(l_L)}{2(k+1)} \right) \right\} \\ & \text{s.t. } l_L \in [0, 4\sqrt{3} - 6] \end{aligned}$$

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 $r_L(\hat{\tau}_L^1, l_L^*, \tau_F^{s1}, l_F^{s1}) = \frac{4}{27}(k+1)$. 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 $l_L \in [0, \frac{1}{2}]$.

Under this supposition, government L 's problem is

$$\begin{aligned} & \max_{(\tau_L, l_L)} \{(\tau_L - l_L) (1 - l_F^{s2})\} \\ & \max_{(\tau_L, l_L)} \left\{ (\tau_L - l_L) \left(\frac{1 - l_Lk - k + \tau_L}{2 - 2k} \right) \right\} \\ & \text{s.t. } l_L \in \left[0, \frac{1}{2} \right] \\ & \tau_L \in [l_L, \hat{\tau}_L^2(l_L)] \end{aligned}$$

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

$$\begin{aligned} & \max_{l_L} \left\{ (\hat{\tau}_L^2(l_L) - l_L) \left(\frac{1 - l_L k - k + \hat{\tau}_L^2(l_L)}{2 - 2k} \right) \right\} \\ & \text{s.t. } l_L \in \left[7 - 4\sqrt{3}, \frac{1}{2} \right] \end{aligned}$$

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 $r_L(\hat{\tau}_L^2, l_L^{**}, \tau_F^{s2}, l_F^{s2}) = \frac{4}{27}(k - 1)$. 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 $(l_L^*, \tau_L^*) = (\frac{8}{9}, \frac{4}{3} + \frac{4}{9}k)$.

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 a 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. (Standard haven) If $\frac{1}{3} < k \leq 1$, then $l_L^* \geq \frac{8}{9}$, $\tau_L^* \in (l_L^*, 2l_L^*)$, and $l_F^* = 0$, $\tau_F^* = \frac{2}{3}(\sqrt{7k^2 - 6k + 3} - k + 3)(k + 1)^{-1} k \in (\frac{3}{4}, \frac{4}{3})$, and $\hat{s}^* = \frac{1}{3}(\sqrt{7k^2 - 6k + 3} - k + 3)(k + 1)^{-1} > \frac{2}{3}$.
- c. (Race to the top) 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] Brennan, G. and J.M. Buchanan, (1980); *The Power To Tax: Analytical Foundations of a Fiscal Constitution*, Cambridge University Press, Cambridge.
- [4] d'Aspremont, C., J. Gabszewicz, and J. Thisse (1979); "On Hotelling's Stability in Competition," *Econometrica*, 17: 1145-1151.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] Hotelling, H. (1929); "Stability in Competition", *Economic Journal*, 39: 41-57.

- [9] 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.
- [10] Niskanen, Jr., W. (1977); *Bureaucracy and Representative Government*, Aldine, Chicago, IL.
- [11] Oates, W.E. and R.M. Schwab (1988); "Economic Competition Among Jurisdictions: Efficiency Enhancing or Distortion Inducing?" *Journal of Public Economics*, 35: 333-354
- [12] 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.
- [13] Taylor, M.S. (2004); "Unbundling the Pollution Haven Hypothesis." Published in *Advances in Economic Analysis and Policy*, the Berkley Electronic Press.
- [14] 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.
- [15] 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.