

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

Valeska Groenert
Ben Zissimos

CESIFO WORKING PAPER NO. 3686

CATEGORY 1: PUBLIC FINANCE

DECEMBER 2011

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

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

Abstract

We show that, in competition between a developed country and a developing country over environmental standards and taxes, the developing country may have a ‘second-mover advantage.’ In our model, firms do not unanimously prefer lower environmental-standard levels. We introduce this feature to an otherwise familiar model of fiscal competition. Four distinct outcomes can be characterized by varying the marginal cost to firms of an environmental externality: (1) the outcome may be efficient; (2) the developing country may be a ‘pollution haven’ - a place to escape excessively high environmental standards in the developed country; (3) the developing country may ‘undercut’ the developed country and attract all firms; (4) the developed country may be a pollution haven.

JEL-Code: H100, H250, H730, H870.

Keywords: environmental standards, fiscal competition, second-mover advantage, tax competition.

Valeska Groenert
MOVE / Department of Economics
Universitat Autònoma de Barcelona
Spain - 08193 Bellaterra (Barcelona)
valeska.groenert@uab.cat

Ben Zissimos
Department of Economics
University of Bath
UK – Bath BA2 7AY
b.c.zissimos@bath.ac.uk

December 2011

We are grateful for the advice we received from Myrna Wooders on this project. Indeed, she was a co-author on an earlier working paper version that had the same title but a different model. We would like to thank Robin Boadway, Sam Bucovetsky, Katherine Cuff, Amrita Dhillon, Will Strange, and three anonymous referees for detailed comments and/or conversations about this paper. Thanks are also due for comments from participants at the Workshop on Stability in Competition, CORE, Louvain la Neuve, a European Science Foundation Workshop, Paris 1, a seminar at the MPI in Bonn, the Social Choice and Welfare Meeting in Istanbul, a Canadian Public Economics Group Meeting at the University of Toronto and seminars at the University of Warwick and Vanderbilt University. Funding from the Center for the Americas at Vanderbilt University and from the Spanish Ministry of Science and Innovation through grant “Consolidated Group-C” ECO2008-04756 is gratefully acknowledged.

1 Introduction

The recent integration of countries in Eastern Europe to the European Union (EU) has provoked renewed concern about the aggressive competition by new members for firms and other mobile factors. For example, although EU accession requirements demand moves towards harmonization of environmental standards and some measures have made it onto statute books, there appears to be widespread skepticism about the actual implementation of such measures. Citing the incentive to keep standards low 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.⁴

The main purpose of this paper is to show that, through competition in ESs and taxes, a developing/transition country may have a ‘second-mover advantage’ over a developed country in attracting firms and extracting rents but not necessarily through the expected channels of low taxes and environmental standards (ESs). 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.

To investigate this concern, our paper develops a model of international competition over ESs and taxes. We will first set out a brief explanation of the model, which will be expanded upon in the paragraphs that follow. There are two countries, one developed and the other developing. There are four firms in the model, one immobile firm in each country and two that are mobile. Mobile firms who locate in a country are required to pay a tax that is used, at least in part, to monitor (and enforce) the ES in that country. Due to monitoring costs, the higher the standard set by a country the more costly it is to implement. The production technology of one of the mobile firms is more polluting than the other. The two immobile firms are more polluting still so, in addition to lowering taxes, governments can compete over ESs to improve the environment in a bid to attract the mobile firms.⁵ The developed country government sets its ES and tax first, followed by the developing country government. Government objectives are purely to maximize revenues. A key parameter in the model is the ‘marginal cost of the pollution externality’ (*mcpe*) which parameterizes how

⁴Andanova (2003) provides further details of environmental policy in Eastern Europe.

⁵We are more used to the suggestion that footloose firms are drawn to the lowest environmental standards. However, there appears to be evidence that this is not universally the case; firms that are internationally mobile are in some cases more productive through their use of more efficient and hence cleaner technologies than domestic firms and hence are attracted to higher environmental standards rather than lower ones (Graham 2000).

a given pollution level affects a firm's costs of production. Each firm chooses its location to maximize profits, taking as given the tax levels and ESs in the two countries as well as the pollution levels of the other firms that locate there.

The issue of competition over ESs and taxes has been raised particularly with respect to the more economically successful 'transition countries' from the former Soviet Union as well as, to a lesser extent, the 'emerging market' developing countries in Asia and the Middle East. The so-called 'Visegrad countries' of the Czech Republic, Hungary, Poland and Slovakia (V4 for short) exemplify the developing and transition countries that we have in mind. These countries are in the midsts of comprehensive governmental reforms, and arguably their governments have a greater degree of flexibility and fewer constitutional and institutional constraints than the long-established democracies in the core of Europe. The same reasoning may also be used to explain the greater flexibility of dictatorships and young democracies further afield. For brevity, throughout the paper we will use 'developing country' as a catch-all term for such countries.⁶ We capture this greater flexibility in policy-making by developing countries, as a result of which they can respond (within the period of a parliament say) to policies adopted by developed countries, through our specification of timing in the game of policy formation. In our two-country model, we will assume that the developed country sets its standard and tax first, followed by the developing country.

In contrast to the past literature (summarized below), we focus on a situation where the tax base is not universally repelled by, nor universally attracted to, ESs. Consider the familiar textbook example where firms need water as an input to production and so must locate around a lake. The assumption that production in one of the mobile firms has a lower propensity to generate pollution than production in the other could be interpreted in one of two ways. In one scenario, both firms are from the same industry but one is newer than the other and, using more efficient technology, pollutes the water in the lake by less per unit of output. In a second scenario, each firm is from a different industry that has a different propensity to generate pollution. In any case we will refer to the firm that generates less

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

pollution as the ‘clean’ firm and the other one as the ‘dirty’ firm. The pollution emitted by the dirty firm reduces the profits of the clean firm more than vice versa. If they both locate in the same country (and around the same lake), an ES that requires all firms to reduce their pollution will improve the profits of the clean firm while the effect on the profits of the dirty one is ambiguous.

Our simple framework yields a surprisingly rich set of equilibrium predictions. There are four possible sorts of outcome which can be characterized as follows. (a) If the *mcpe* is low and the mobile firms are similar in their pollution levels then fiscal competition leads to an efficient equilibrium outcome (as in Brennan and Buchanan’s 1980 model of tax competition). (b) If the *mcpe* is low but the mobile firms differ to a greater degree in their pollution levels then the developed country may set its ES inefficiently high, in which case the developing country becomes a pollution haven; a place where the dirty firm locates in order to escape the high ES set in the developed country. (c) If the *mcpe* is high and the mobile firms are similar in their pollution levels then the developing country is able to undercut the developed country, with both firms locating in the developing country. (d) If the *mcpe* is high and the mobile firms differ to a greater degree in their pollution levels then the developed country becomes the pollution haven, where the dirty firm locates to escape a (not necessarily inefficiently) high ES set by the developing country. It is especially interesting that inefficiently high standards can arise in equilibrium, either in the developed country (as in b) or in the developing country (as in c and d) purely through strategic interaction between governments in their competition for firms and not as a result of attempts by governments to reduce emissions on behalf of consumers/citizens.⁷

The key strategic consideration that drives our results is that the developed country wants to bring about an outcome in which the developing country is prepared to ‘share’ the mobile firms rather than undercut the developed country and attract them both. To do so, the developed country must put the developing country in a situation where it can earn higher rents by sharing firms than by undercutting. Intuitively, sharing is particularly

⁷The issue of governments setting environmental standards too high, in order to dissuade a noxious production facility from locating in their jurisdictions, is referred to as ‘not in my back yard’ or NIMBY, and brings about a ‘race to the top’. NIMBY has been studied by Levinson (1999) among others. Conversely, Kempf and Rossignol (2007) study a situation where voters object to a higher burden from environmental policy being placed on the industry where they work relative to other industries if this adversely affects their welfare, say through employment. In our framework governments do not repel firms in response to environmental concerns by consumers nor other parties who may be harmed by the hosting of firms. Yet we can get overprovision or underprovision of the ES through a quite different set of interactions.

beneficial for the developing country if it attracts the clean firm and leaves the dirty firm to the developed country. There are two reasons to think this. The first is simply that the clean firm is cheaper to monitor since it has a lower incentive to deviate from the ES. The second more interesting effect is that the clean firm is less elastic in its location decision than the dirty firm; to avoid the relatively large externality exerted by the dirty firm it prefers to locate in the other country, and this creates extra rent that the government can extract from the clean firm through higher taxation.⁸

From this perspective, the most surprising equilibrium outcome is actually case (b) where the developing country becomes a pollution haven, setting the minimum ES and attracting the dirty firm. In this case the developed country can make sharing firms attractive for the developing country by setting its ES at a high level and attracting the clean firm; at a low *mcp_e*, this high ES level makes the developed country unattractive to the relatively noxious dirty firm. Consequently, the developing country can set the minimum ES and a relatively high tax and still attract it.

Standard international trade theory provides a different rationale for the existence of pollution havens, based on the argument that lax ESs are a source of comparative advantage since they lower the opportunity cost of pollution. However, the empirical evidence for this motivation for pollution havens is mixed (see Antweiler, Copeland and Taylor 2001, Taylor 2004, and Ederington, Levinson and Minier 2004). Therefore, our case (b) is helpful in that it presents an alternative strategic motivation for the existence of pollution havens in developing countries based on their second-mover advantage.

An interesting feature of equilibrium in our model is that the developing country practices limit-taxing by setting a tax just low enough to attract the firm(s) that it wishes to attract. The developing country's power to extract rents through limit-taxing arises solely through its second-mover advantage. This contrasts with the prior literature, exemplified by Markusen, Morey and Olewiler (1995) and Baldwin and Krugman (2004), in which it is the developed country that has a first-mover advantage derived from its power to act as a limit-pricing monopolist because at the outset firms are located there and have an 'attachment to home.'

⁸One objection might be that developing countries typically set relatively low taxes as well as ESs whereas in our model the tax set by developing countries can be relatively high. A simple 'fix' to our model which would enable the developed country to set higher taxes would be to initially locate the mobile firms in the developed country and give them an attachment to home. This would enable us to control relative taxation across countries by varying the 'attachment to home' parameter. The reason we didn't adopt this fix was because, while the outcome would have been more realistic, the workings of the model would have been made less transparent. Other possible modifications to the model are discussed in the concluding section.

Broadly, the prior literature on interjurisdictional competition over ESs and taxes can be categorized into two areas. The first area, following Tiebout (1956), focuses on situations where competition among independent governments is like competition among firms and enhances efficiency. The ‘Tiebout assumption’ is that all firms benefit to differing degrees from a clean environment and sort themselves efficiently into jurisdictions each of which enforces an ES that is appropriate for its members. The second area concerns the presence of a policy-failure that allows or induces governments to set taxes on mobile capital, as in the literature on fiscal federalism and ‘standard tax competition’ associated with Oates (1972), Wilson (1986) and Zodrow and Mieszkowski (1986). Capital is indifferent to the imposition of an ES, but is repelled if burdened with having to pay for the ES. In these situations local jurisdictions, while competing for mobile capital, at the same time tax that capital to protect the environment. (In this literature, the terms ‘environmental standard’ and ‘environmental regulation’ are used interchangeably; see Wilson 1996 and Levinson 2003 for surveys.) Our model combines features of models from papers in both areas: on the one hand firms differ in the degree to which they benefit from a clean environment and 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 far as we are aware, the situation that we examine here in which firms inflict pollution externalities on each other has not previously been studied in the context of fiscal competition. And the issue of developing country second-mover advantage has not formally been motivated. The model is of course highly stylized. Nevertheless, the results seem intuitively plausible and may be indicative of more general strategic interactions over taxes and environmental policy that have aroused substantive concerns in the policy debate.⁹

The remainder of this paper is organized as follows. Section 2 introduces the model. Section 3 solves for the efficient allocation. Section 4 defines strategies and the subgame perfect equilibrium and then characterizes equilibrium in terms of the four cases outlined above. Conclusions are drawn in Section 5.

⁹In Groenert, Wooders and Zissimos (2009) we delivered a very similar set of results using a different model based on a continuum of firms. This suggests that the results we deliver here are more general than might be suggested by the specific set-up of our model. The reason for moving to the present simpler specification is that as a result the details of the key strategic consideration discussed above, which we regard to be a contribution of the paper, could be brought out more sharply. In addition, the pollution externality could be modeled in a more satisfactory way in a framework with just four firms than with firms on a continuum.

2 The Model

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

2.1 Firms

Each firm is able to sell its single unit at price p and has a fixed private per-unit production cost, γ .¹⁰ The tax levied on the firm is τ_L if it locates in L and τ_F if it locates in F . Let the variables $l_L, l_F \in [0, 1]$ denote the ES levels in L and F respectively. Let $s_j \in [0, 1]$ denote the per-unit-of-output pollution level a given firm j emits in the absence of any ES. There is an immobile ‘home’ firm in each country, h , for which $s_h = 1$. In addition there are two mobile firms: the clean firm, c , with pollution level s_c ; and the dirty firm, d , with pollution level s_d . Let $s_c \leq s_d < 1$, and normalize so that $s_c = 0$.

Firms that are located in the same country impose a nonpositive externality on each other. In the absence of any ES, the externality that a firm s' imposes on a firm s depends on the pollution level of each firm as well as on a parameter $k \in \mathbb{R}_+$ and is given as follows:

$$ks'(1 - s).$$

The dirtier the firm s' the larger the externality imposed on firm s . Moreover, the cleaner is a firm’s technology, the more damage pollution by another firm does to it.¹¹ The overall

¹⁰Prices and costs could be made to vary across firms without affecting the results.

¹¹This assumption captures the idea that one firm’s pollution contaminates the environmental resource that another firm uses in its production. If, for example, a firm uses water as a cleaning agent in its production process, and if that water has been polluted, then the firm must clean the water before using it. The underlying assumption would then be that dirtier water is costlier to clean. Conversely, the production process of a dirtier firm is less affected by the pollution of other firms. This assumption is consistent with a wide variety of relatively dirty production processes that do not require a high level of purity in the water that they use. But it does exclude, for example, the special case of a nuclear power-plant that needs pure water in its production process but whose pollution is extremely costly for other firms since it turns the water radioactive.

impact of the externality is captured by the parameter k ; this is what we refer to as *mcpe*. The larger is k the larger the negative impact of the pollution by other firms on a firm's profit. If the ES level in country i is set at l_i and the firm locates in country i , the impact of the externality imposed by s' on s is reduced to

$$k(1 - l_i) s'(1 - s).$$

Thus an ES level l requires that a proportion l of the cost of pollution be internalized.¹² If there are several polluting firms in a country the negative externalities add up. Let M_i denote the set of firms that are located in country $i \in \{L, F\}$. Then the pollution externality suffered by firm j , $j \in \{c, d, h\}$, can be expressed as

$$k(1 - l_i) E_i^j(1 - s_j),$$

where $E_i^j = \sum_{m \in M_i \setminus j} s_m$.¹³

Abiding by a certain ES level is costly for firms. Moreover, the dirtier the firm is, the costlier it is for that firm to abide by the ES. This property is captured in the profit function by $s_j l_i$. Thus the profit of firm j that locates in country i is calculated as follows:

$$\pi(j) = p - \gamma - \tau_i - s_j l_i - k(1 - l_i) E_i^j(1 - s_j), \quad i \in \{L, F\}.$$

To focus the analysis on location decisions, it will be assumed throughout that p is sufficiently high to ensure that all firms make nonnegative profits wherever they locate. A firm locates in the country where its profits are maximized. If a firm's profits are the same in both countries we assume that it locates in F .¹⁴

Firm j may prefer one country, say F , in terms of the tax that it sets; $\tau_F < \tau_L$. But if the environmental standards in L and F and the location of the other firms are such that $s_j l_F + k(1 - l_F) E_F^j(1 - s_j) > s_j l_L + k(1 - l_L) E_L^j(1 - s_j)$, firm j may nonetheless prefer to locate in L . Notice that, all else equal, the clean firm always prefers a higher ES level since its production technology is already as clean as possible. Moreover, unless the ES is set at

¹²It is standard to regard an ES as a ceiling on pollution emissions, such as an upper bound on the density of a particulate in waste water. Our framework fits with this interpretation under the assumption that any ES level $l \in (0, 1]$ imposes a binding constraint on firms d and h but not on firm c and that firm h is required to reduce its emissions by an amount that is proportionately greater than firm d is.

¹³Say for example that the clean firm but not the dirty firm locates in country i . Then for the home firm located in that country, $E_i^h = s_c$, while for the clean firm $E_i^c = s_h$.

¹⁴This assumption ensures existence of a rent-maximizing strategy among F 's strategies that attracts both mobile firms. The same result could be obtained by introducing arbitrarily small indivisibilities in the strategy sets. For ease of exposition, we chose not to do that.

the highest level, the clean firm is better off if the dirty firm locates in the other country. In contrast, it is costly for the dirty firm to abide by the ES and therefore it prefers higher ES levels only if k is sufficiently high. Since $s_c = 0$, the dirty firm is indifferent about the location of the clean firm and hence is only concerned about the level of the ES because of its own cost of compliance and the limit that the ES imposes on the externality generated by the home firm.

2.2 Governments

Rents are given by tax revenues minus the cost of ES-setting. We assume that a government can vary a tax only for the mobile firms that locate in its country, whereas the home firm pays a tax or receives a subsidy that is fixed.¹⁵ Besides a tax, each government i chooses an ES level, l_i , which can take on any value in the unit interval. The pollution levels of firms located in country i sum to $E_i = \sum_{j \in M_i} s_j$. Then, since the government of i must enforce the ES on all its firms, its cost of monitoring an ES level, $l_i \in [0, 1]$, is $E_i l_i$. Thus the cost of monitoring a given ES is assumed to be proportional to the level of the ES and it is more costly to monitor the ES for dirtier firms (who have larger incentives to evade the ES).¹⁶ Governments move sequentially. Government F observes l_L and τ_L and chooses l_F and τ_F to maximize its rents.¹⁷ Here is the rent function for government F ; the rent function for L is symmetric:

¹⁵The interpretation is that there is a ceiling on the tax (or equivalently, a floor on the subsidy) for the home firm. This ceiling/floor could be an outcome of a strong lobby of the home firm. Cases in point are the oil, gas, and coal mining industries whose domestic firms continue to receive subsidies throughout the EU (Euractiv 2010 and New York Times 2010). In the absence of this assumption, the incentive to use the tax to extract all profits from the (immobile) home firm may have an overbearing influence on government incentives.

¹⁶We would not expect our results to change qualitatively if the costs of ES setting were strictly convex instead of linear.

For greater tractability, we have adopted a reduced-form approach to incorporating the ES monitoring costs. The following structural model of ES monitoring would generate the same cost structure. Suppose that for each ES level of l a firm j has to decide on the level of pollution reduction $\rho \in [0, 1]$ it undertakes at cost ρs_j . The firm has to pay a fine of $\text{Max}\{0, f(l - \rho)\}$ if the government finds that it has violated the ES. To enforce $\rho = l$, the government has to choose a probability of monitoring the firm of at least $\text{Pr} = s_j/f$. Thus, firms that have a bigger incentive to pollute need closer and hence more costly monitoring.

¹⁷For any of the following variations in the sequence of moves in our model, equilibrium in pure strategies does not exist: 1) Governments set ESs and taxes simultaneously; 2) L sets an ES level, then, after learning L 's choice, F sets its ES level, and then, both governments simultaneously set taxes; 3) governments first simultaneously set ESs and then simultaneously set taxes.

$$r_F(l_F, \tau_F; l_L, \tau_L) = \begin{cases} \tau_F - l_F & M_F = \{c, h\} & \text{(A)} \\ \tau_F - l_F(1 + s_d) & M_F = \{d, h\} & \text{(B)} \\ 2\tau_F - l_F(1 + s_d) & M_F = \{c, d, h\} & \text{(C)} \\ -l_F & M_F = \{h\} & \text{(D)} \end{cases} \text{ if } .$$

The rent function does not include the tax/subsidy for the home firm since it does not play a role in our analysis. To obtain clear-cut solutions for firm locations, we assume the two following tie-breaking rules in cases where governments have multiple best-responses: (1) If indifferent, F chooses sharing strategies within (A) or (B) over the undercutting strategies (C), and sharing strategies that attract the clean firm (A) over sharing strategies that attract the dirty firm (B); (2) If L has multiple best responses it chooses the one with the lowest ES level and so does F if indifferent among strategies within A, B, C, or D. An alternative to the first assumption would be to introduce indivisibilities into the strategy space, and the second assumption plays a role only at single points in the parameter space of k and s_d .

3 Efficiency

We define the efficient outcome to be the one that a social planner would choose if he could freely choose ES levels (while still having to pay the costs of monitoring) and the allocation of mobile firms to countries, but not the allocation of the immobile firms. Therefore the efficient outcome need not respect the incentives of mobile firms to change location.¹⁸ The planner is assumed to maximize the aggregate surplus realized by firms plus the governments' rents. In this section, we refer to the two countries with indices 1 and 2 since it does not matter for the efficiency of the allocation which one is F and which one is L . An allocation consists of two ES levels and an assignment of mobile firms to countries, denoted by (l_1, l_2, M_1) . There are only two types of allocations of firms to consider: 1) The mobile firms are separated; 2) both mobile firms are allocated to the same country.

(1) Without loss of generality, let $M_1 = \{d, h\}$. To derive the optimal ES level in country 1 solve

$$\min_{l_1} \{(l_1 + s_d l_1) + k(1 - l_1)(1 - s_d) + (1 + s_d)l_1\}.$$

The first term is the cost of abiding by the ES incurred by the home firm and the dirty firm in country 1, the second term is the externality that the home firm imposes on the dirty firm,

¹⁸The efficient outcome does not change if the planner has to respect mobile firms' incentives to relocate. This is because the planner can always set the efficient locating incentives through taxes (which are welfare neutral).

and the third term is the ES monitoring cost that the planner incurs. Solving this yields a critical level of $k = \frac{2(1+s_d)}{1-s_d}$ such that: for $k \leq \frac{2(1+s_d)}{1-s_d}$ the planner chooses the minimum ES level, $l_1 = 0$; for $k > \frac{2(1+s_d)}{1-s_d}$ the planner chooses the maximum ES level, $l_1 = 1$.¹⁹ To derive the optimal ES level in country 2, where the clean firm is located, solve

$$\min_{l_2} \{l_2 + k(1 - l_2) + l_2\}.$$

The interpretation of the three terms is the same as in the previous equation but for country 2. Solving this for $k \leq 2$ yields $l_2 = 0$ and for $k > 2$ yields $l_2 = 1$.

(2) Without loss of generality, let $M_1 = \{h\}$. Thus the optimal ES in country 1 is $l_1 = 0$. To derive the optimal ES level in country 2, solve

$$\min_{l_2} \{(l_2 + s_d l_2) + (k(1 - l_2)(1 + s_d) + k(1 - l_2)(1 - s_d)) + (1 + s_d)l_2\}.$$

Solving this for $k \leq 1 + s_d$ yields $l_2 = 0$ and for $k > 1 + s_d$ yields $l_2 = 1$.

Notice that in each of the cases just solved, there exists a critical level of k such that the optimal ES level is set at minimum if k is below that level and at maximum for higher levels of k . This critical level of k is smallest in allocation (2), in which the planner allocates both mobile firms to the same country (so that the benefit of the ES is largest) and is largest if only the dirty firm is allocated to a country; that is $1 + s_d < 2 \leq \frac{2(1+s_d)}{1-s_d}$.

To decide between separating firms versus locating them in the same country, we compare the costs resulting from the optimal ES levels in (1) and (2). Results differ depending on the level of s_d . Table 1, in which (l_1^e, l_2^e, M_1^e) denotes the efficient allocation, summarizes our results; Figure 1 illustrates them.

Table 1: Efficient Outcome

k	l_1^e	l_2^e	M_1^e
$k \leq 2$	0	0	$\{d, h\}$
$k \in \left(2, \frac{2s_d}{1-s_d}\right]$ (non-empty only if $s_d > \frac{1}{2}$)	0	1	$\{d, h\}$
$k > \text{Max} \left\{2, \frac{2s_d}{1-s_d}\right\}$	0	1	$\{h\}$

For $k \leq 2$, separation of firms with both countries setting the minimum ES is optimal. For $k > 2$, it is optimal to set the minimum ES in one country and the maximum in the other. If $k \in \left(2, \frac{2s_d}{1-s_d}\right]$, separation is optimal, with the clean firm located in the country with the

¹⁹We select the lowest ES level if there is more than one efficient one.

maximum ES, while if $k > \text{Max} \left\{ 2, \frac{2s_d}{1-s_d} \right\}$ (or equivalently, $s_d < \hat{s}_d \equiv \frac{k}{k+2}$ and $k > 2$), both mobile firms should be located there. In Figure 1, (x, y) is short for $l_1 = x$ and $l_2 = y$. We say that firms are: ‘Separate’ when one mobile firm is located in each country and, if one country sets the maximum ES, the clean firm is located there; ‘Together’ when both mobile firms are located in the country that sets the maximum ES.

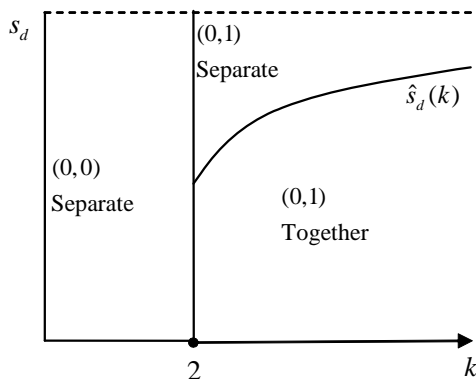


Figure 1

The efficient ES levels are weakly increasing in k in the sense that for $k \leq 2$ the planner sets the minimum ES level in both countries while for $k > 2$ he sets the ES at its maximum level in one country, maintaining the minimum ES level in the other. If s_d is large relative to k , in which case abiding by the ES is relatively costly for the dirty firm and monitoring the dirty firm is as well, the dirty firm is most efficiently allocated to the country with the minimum ES. This efficient solution will serve as a benchmark against which to compare the equilibrium outcome.

4 Competition over Environmental Standards and Taxes

In this section, our approach will be to first define equilibrium and then state our main theorem in which equilibrium is characterized. After that, we will provide some intuition for our results and sketch their derivation. The full proof is presented in the Appendix.

As mentioned above, ES provision and tax setting are modeled as a two-stage game. Government L sets its ES level and tax and then, observing L 's choices, government F sets its ES level and tax. Taking government policies as given, firms then make location decisions to maximize profits. As usual, a *subgame perfect Nash equilibrium* is a strategy

profile with the property that the governments' strategies constitute a Nash equilibrium in every subgame of the game.

A strategy for government L is a pair consisting of an ES level and a tax. Formally, the set of strategies is $S_L = (l_L, \tau_L) \in [0, 1] \times \mathbb{R}_+$. A strategy for government F is a mapping that assigns a pair, consisting of an ES level and a tax, to each possible strategy choice made by government L in the first stage of the game. Formally, this mapping is described by $f : S_L \rightarrow [0, 1] \times \mathbb{R}_+$ where $f(l_L, \tau_L) = (l_F, \tau_F)$. Let \mathcal{F} be the set that contains all such mappings. The set of strategies for government F is $S_F = \mathcal{F}$.

We are interested in the pure strategy subgame perfect Nash equilibrium of the game, which can be viewed as a Stackelberg game.²⁰

Definition 1. *A pure strategy subgame perfect Nash equilibrium in taxes and ES levels is a pair of strategies $((l_L^*, \tau_L^*), f^*)$ such that: (1) $(l_L^*, \tau_L^*) \in S_L$ is a best response to f^* ; (2) $f^* \in S_F$ and $f^*(l_L, \tau_L)$ is a best response to (l_L, τ_L) for all $(l_L, \tau_L) \in S_L$.*

With the structure of the model in place and equilibrium defined, we are now ready to state our main theorem which characterizes equilibrium. For use in the theorem and as illustrated in the following figure, the function $\tilde{s}_d(k)$ is the boundary between cases a and b of the theorem, and the function $\bar{s}_d(k)$ is the boundary between case (c) and cases (a) and (d).²¹ Let M_F^* denote the set of firms that locate in F in equilibrium.

Theorem 1. *The subgame perfect equilibrium is as follows.*

- a. *(Efficient outcome) If $s_d \leq \tilde{s}_d(k)$ and $k \leq \text{Min} \left\{ 1, \frac{3s_d+1}{2(1-s_d)} \right\}$, both L and F set the minimum ES level. Firms separate with the cleaner firm locating in F . Specifically, it holds that $l_L^* = 0$, $l_F^* = 0$, $\tau_F^* = \tau_L^* + ks_d$, and $M_F^* = \{c, h\}$. The outcome is efficient.*
- b. *(Pollution haven in F) If $s_d > \tilde{s}_d(k)$, the differentiation in ES levels between the two countries is high; L sets the maximum ES level and F sets the minimum ES level. Firms separate with the dirty firm locating in F . Specifically, it holds that $l_L^* = 1$,*

²⁰It 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.

²¹The exact expressions for these functions are as follows:

$$\tilde{s}_d(k) = \begin{cases} 1 - k & \text{if } k \leq \frac{1}{2} \\ \frac{2k}{2-2k^2+k} & \text{if } k > \frac{1}{2} \end{cases}; \bar{s}_d(k) = \begin{cases} \frac{2k-1}{3(k+1)} & \text{if } k \leq 1 \\ \frac{k}{3(k+1)} & \text{if } k > 1 \end{cases}$$

$l_F^* = 0$, $\tau_F^* = \tau_L^* + (s_d - k + ks_d)$, and $M_F^* = \{d, h\}$. There is overprovision of the ES in L.

- c. (*F* undercuts *L*) If $s_d < \bar{s}_d(k)$ (which implies $k > \frac{1}{2}$), the differentiation in ES levels between the two countries is high; *F* sets the maximum ES level and *L* sets the minimum ES level. Both firms locate together in *F*. Specifically, it holds that $l_L^* = 0, l_F^* = 1$, and $\tau_F^* = \tau_L^* - (s_d - k + ks_d)$, and $M_F^* = \{c, d, h\}$. For $k \leq 2$ there is overprovision of the ES and for $k > 2$ the outcome is efficient.
- d. (*Pollution haven in L*) If $s_d \geq \bar{s}_d(k)$ and $k > 1$, the differentiation in ES levels between the two countries is high; *F* sets the maximum ES level and *L* sets the minimum ES level. Firms separate with the clean firm locating in *F*. Specifically, it holds that $l_L^* = 0, l_F^* = 1$, $\tau_F^* = \tau_L^* + k(1 + s_d)$, and $M_F^* = \{c, h\}$. For $k \leq 2$ there is overprovision of the ES, and for $k > 2$ there is underprovision for $s_d < \hat{s}_d(k)$ and the outcome is efficient for $s_d \geq \hat{s}_d(k)$.

There is always a second-mover advantage, that is rents (and taxes) are higher for *F* than for *L*. Except for if *F* undercuts *L*, both governments make positive rents.

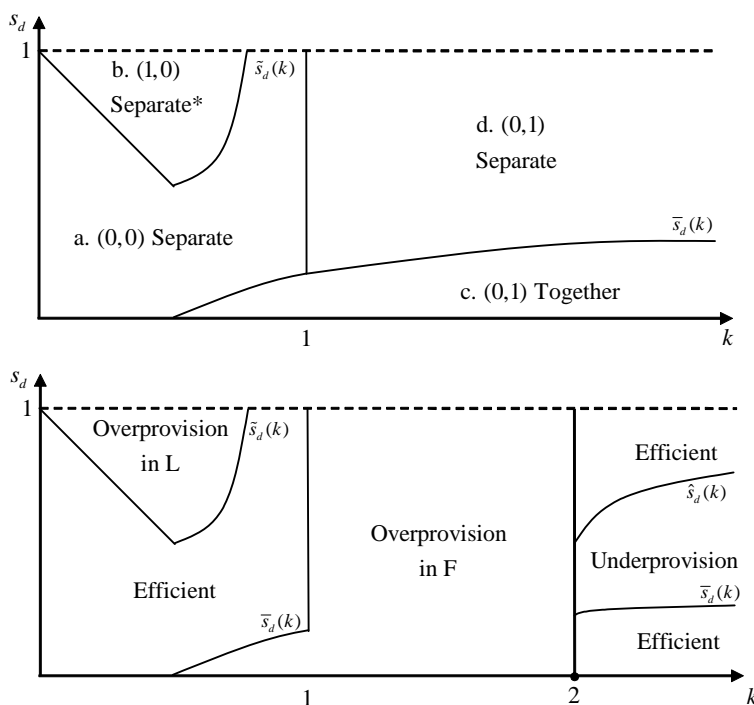


Figure 2

(Note that the labeling of the cases as a to d in the theorem is unrelated to the follower's four strategies, labeled above as A to D .) All results depend on the two key parameters k and s_d as Figure 2 illustrates. The upper panel shows the four (qualitatively different) types of equilibrium; the lower panel shows whether, compared to the efficient outcome, there is too much ('Overprovision') or too little ('Underprovision') of the ES. In the figure, 'Separate' refers to the case where the clean firm locates in F and the dirty firm locates in L and 'Separate*' refers to the reverse case. Subgame perfect ES levels and the locations of firms differ considerably across the four regions of k and s_d . In the lower panel, the function $\hat{s}_d(k)$ separates efficiency from underprovision for $k > 2$.²²

The key aspect that drives the results is that L wants to induce F to share firms rather than undercut. To do so, L has to make it attractive for F to share the firms by giving F an incentive to incur relatively low costs of monitoring its ES and/or setting its own policy in such a way that it does not attract the firm from which F would reap higher rents.

In case (a), where k is small and s_d is below the threshold value $\tilde{s}_d(k)$, the externality is not very important for either of the firms and setting the minimum ES level is a dominant sharing strategy for F . However, even for relatively low (but still positive) s_d the clean firm tries to avoid locating in the same country as the dirty firm, and this makes its location decision less elastic with respect to the tax than that of the dirty firm. This relative inelasticity generates rent that can be extracted from the firm by the government where it locates. To avoid being undercut L attracts the dirty firm, leaving the clean firm and the associated higher rent to F .

In case (b), where k is also small but s_d is relatively large, L makes sharing an attractive strategy for F by setting the maximum ES level and attracting the clean firm; this maximum ES level makes L unattractive to the dirty firm, thereby allowing F to set a high tax and extract relatively high tax revenue from it. Since the efficient outcome for this parameter range is that both governments set the minimum ES level, there is overprovision in L .

In cases (c) and (d), in which k is bounded from below, attracting only the clean firm dominates attracting only the dirty firm for F because the clean firm's willingness to pay not to be in the same country as the dirty firm is high relative to the cost of monitoring. Since k is large, the best way to attract the clean firm is to set the maximum ES level. Given this, setting the minimum ES level is best for L since this is cheap and further deters the clean

²²The exact expression for $\hat{s}_d(k)$ is $\hat{s}_d(k) = \begin{cases} 0 & \text{if } k < 2 \\ \frac{k}{2+k} & \text{if } k \geq 2 \end{cases}$.

firm from L . If, as in case (d), s_d is above $\bar{s}_d(k)$, L is successful at inducing F to attract only the clean firm. Relative to the efficient outcome, this means that F sets an inefficiently high standard if $k \leq 2$. If $k > 2$ there is ‘underprovision’ of the ES (in the sense that the dirty firm goes unmonitored and would be more efficiently allocated to F , where the ES is set at the maximum level) for s_d below $\hat{s}_d(k)$, and the outcome is efficient otherwise. In contrast, if s_d is low, as in case (c), L does not succeed in making sharing firms sufficiently profitable for F . Undercutting dominates attracting only the clean firm as the two firms are similar in their preferences and the maximum ES level eliminates the pollution externality between them. Relative to the efficient outcome, this means that F sets an inefficiently high ES level for $k \leq 2$, while the outcome is efficient for higher levels of k .

The common characteristic of equilibrium across all levels of k and s_d is that there is a second-mover advantage in that F is always able to extract more rents than L . Both governments make positive rents, except for L in case (c) where s_d is relatively low and k relatively large. The ability to extract rents arises as a result of the monopolistic power that each government has over location within its country. Each firm must locate in one country or the other in order to produce, and the government of the country where it does locate is able to exploit its resultant power when setting taxes. In the following two subsections we discuss in more detail the strategic considerations of the two countries.

4.1 The Developing Country

The government of the developing country, F , given a strategy (l_L, τ_L) by L , chooses between four basic strategies: (A) attract the clean firm but not the dirty one, (B) attract the dirty firm but not the clean one, (C) attract both mobile firms, and (D) do not attract either of the mobile firms. We show in the Appendix that F never chooses a strategy in D.²³ To find the best-response strategy, we first determine the optimal strategy within each of the remaining three types of strategies and then compare the payoffs from each.

A) To attract only the clean firm, F must choose (l_F, τ_F) so that if the clean firm locates in F and the dirty firm locates in L , neither firm has an incentive to move to the other

²³The brief argument is as follows. By setting the minimum ES level and a tax sufficiently high to deter both mobile firms, L can always ensure nonnegative rents. Therefore, in equilibrium L sets a tax that covers its costs. For any such strategy, it can be shown that F has a best response that earns it positive rents. Every strategy in D leads to at most zero rents and is therefore not chosen by F in equilibrium.

country. Thus we need

$$\begin{aligned}\tau_F + k(1 - l_F) &\leq \tau_L + k(1 - l_L)(1 + s_d), \text{ and} \\ \tau_L + s_d l_L + k(1 - l_L)(1 - s_d) &< \tau_F + s_d l_F + k(1 - l_F)(1 - s_d),\end{aligned}$$

or

$$\begin{aligned}\tau_F &\in (A', A''] \\ &\equiv (\tau_L + (l_L - l_F)(s_d - k + ks_d), \tau_L + k(l_F + s_d - l_L(1 + s_d))].\end{aligned}$$

B) Similarly, to attract only the dirty firm, we need

$$\begin{aligned}\tau_F &\in (B', B''] \\ &\equiv (\tau_L + k(l_F - l_L - s_d + l_F s_d), \tau_L + (l_L - l_F)(s_d - k + ks_d)).\end{aligned}$$

C) And to attract both mobile firms, we need

$$\begin{aligned}\tau_F &\leq \text{Min}\{C', C''\} \\ &\equiv \text{Min}\{\tau_L + k(l_F - l_L - s_d + l_F s_d), \tau_L + (l_L - l_F)(s_d - k + ks_d)\}.\end{aligned}$$

Define $t_A \equiv A'$, $t_B \equiv B'$ ($= A' = C''$) and $t_C \equiv C'$ ($= B'$), which are all linear functions of l_F . Table 2 lists the intercepts, α_j , and slopes, β_j , of t_j , for $j \in \{A, B, C\}$. These have natural interpretations that will be discussed below.

Table 2: Boundaries of Strategy Sets

j	α_j	β_j
A	$\tau_L - kl_L + ks_d(1 - l_L)$	k
B	$\tau_L - kl_L + l_L s_d(1 + k)$	$k - s_d(1 + k)$
C	$\tau_L - kl_L - ks_d$	$k + ks_d$

It is easy to see that

$$\begin{aligned}\alpha_C &\leq \text{Min}\{\alpha_A, \alpha_B\}, \text{ where } \alpha_A < \alpha_B \Leftrightarrow l_L > \frac{k}{1 + 2k}, \text{ and} \\ \beta_B &\leq \beta_A \leq \beta_C.\end{aligned}$$

Figure 3 depicts the different strategy sets for $l_L \leq \frac{k}{1+2k}$ (left panel) and $l_L > \frac{k}{1+2k}$ (right panel). To find the optimal strategy within each subset F chooses a standard level and the

matching maximum tax in the strategy subset (where, in the figure, this is the tax at the upper boundary of the respective set). From F 's rent function we have:

$$r_F(l_F, \tau_F; l_L, \tau_L) = \begin{cases} \tau_F - l_F & \text{if } \tau_F \in (t_B, t_A] & \text{(A)} \\ \tau_F - l_F(1 + s_d) & \text{if } \tau_F \in (t_C, t_B] & \text{(B)} \\ 2\tau_F - l_F(1 + s_d) & \text{if } \tau_F \leq \text{Min}\{t_C, t_B\} & \text{(C)} \\ -l_F & \text{else} & \text{(D)} \end{cases} .$$

Two useful observations about the optimal strategy can be made right away from inspection of the rent function and Figure 3. First, since monitoring an ES level is more costly for F if it attracts the dirty firm (B) instead of the clean firm (A), and since for $l_L \leq \frac{k}{1+2k}$, the maximum tax F can set in A is always at least as high as the one in B ($\alpha_A \geq \alpha_B$ and $\beta_A \geq \beta_B$), all strategies in B are dominated. Second, if the maximum tax F can set decreases in the standard, raising the standard can only decrease F 's rent. The (l_F, τ_F) combinations that are potential best responses, taking into account the two observations just made, are indicated by bold lines in Figure 3.²⁴

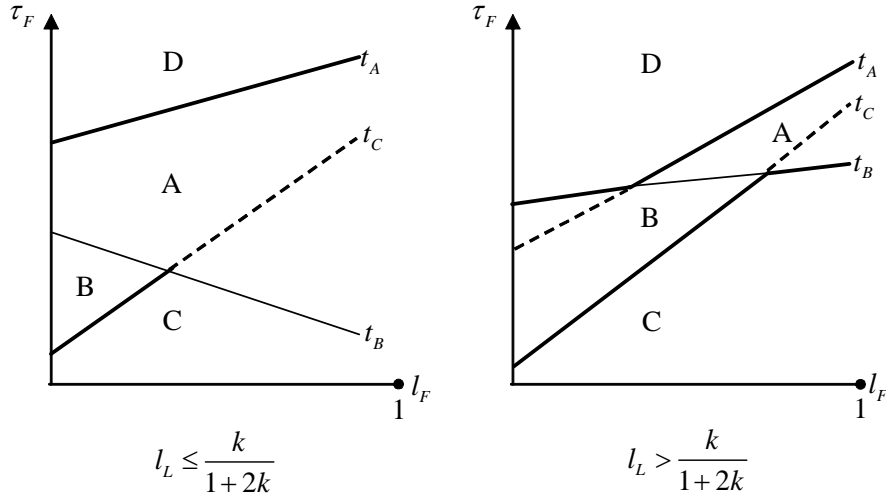


Figure 3

Taking strategies in B as an example, let us consider the trade-off F faces when choosing an ES level. For each of the three kinds of strategies, this trade-off is captured by the intercepts α_j and slopes β_j , for $j \in \{A, B, C\}$. If F wishes to attract the dirty firm and set an ES of zero, the maximum tax it can set is $\alpha_B = \tau_L + s_d l_L - k l_L(1 - s_d)$. This expression is composed of the tax in L plus the cost of abiding by the ES that the dirty firm faces in

²⁴Note that although t_C lies below t_A and (for some l_F) t_B , setting t_C can potentially yield a higher rent since it attracts both mobile firms instead of only one.

L , minus the dirty firm's cost reduction in L due to the ES there. If F sets a positive ES level of l_F instead, it has to modify the tax by $l_F\beta_B = l_F(k(1-s_d) - s_d)$; the dirty firm's cost reduction due to the ES, $kl_F(1-s_d)$, minus its cost of abiding by the standard, l_Fs_d . Similarly, in A , the government of F has to take into account the clean firm's incentive to locate in L instead. In C , only one of the mobile firms constitute a binding constraint on the maximum tax F can set: In the figure, where for low l_F the upper boundary of C borders with B , it is the clean firm whose incentives matter, because it has a strong incentive to escape the externality from the dirty firm; for large l_F the upper boundary of C borders with A and it is the dirty firm whose incentives matter, because it has a strong incentive to escape the high cost of abiding by the ES.

To find the rent maximizing ES levels in each of the three cases, F compares the marginal willingness to pay for the standard of the firms it wishes to attract with its own marginal cost of monitoring the standard, which is 1 in case A and $(1+s_d)$ in B and C . Since a mobile firm's marginal willingness to pay increases as the externality becomes more important, the optimal ES level for each of the three cases is weakly increasing in k . In contrast, the optimal ES level is weakly decreasing in s_d , the degree to which the mobile firms differ.

The last step to find F 's best response is to compare rents from the optimal strategies in each of the three cases. Since, for $l_L \leq \frac{k}{1+2k}$, A dominates B it only remains to compare A and C . Since in C , government F collects the tax twice and the maximum tax it can charge increases with τ_L by the same rate in both cases (as can be seen from the intercepts α_A and α_C in Table 2), eventually with sufficiently large τ_L , the rent-maximizing strategy in C yields higher rents than the rent-maximizing strategy in A . Thus there is a critical level of L 's tax, $\hat{\tau}_L(l_L; k, s_d)$, which we call the *sharing tax limit*, such that F attracts the clean firm up to that tax (i.e. it 'shares' firms with L) and undercuts if L 's tax exceeds $\hat{\tau}_L(l_L; k, s_d)$. (This is not immediate from Figure 3 but will be illustrated in the next subsection). Similarly, if $l_L > \frac{k}{1+2k}$, there is such a sharing tax limit. Up to this tax F shares firms, by attracting only the clean firm for lower levels of l_L and only the dirty firm for higher levels of l_L .

4.2 The Developed Country

The government of the developed country, L , can only make positive rents if it can induce F to share firms. Thus L chooses a tax that does not exceed the sharing tax limit. Since rents are increasing in taxes L will choose $\tau_L = \hat{\tau}_L(l_L; k, s_d)$. It only remains to determine the

optimal ES level. In choosing an ES level, L takes into account two aspects. First, L tries to find an ES level at which, compared to undercutting, sharing is particularly beneficial for F so that the sharing tax limit is high. Second, L takes into account the cost of monitoring the ES. Figure 4 shows the sharing tax limit (the solid line) and the cost of monitoring the ES (the dashed line) at four different combinations of s_d and k , one for each of the four different equilibrium outcomes. For L the optimal level of the ES is where the difference between the sharing tax limit and monitoring costs is maximized. If $\hat{\tau}_L(l_L; k, s_d)$ is always at or below its monitoring cost L cannot make positive rents (as in case c).

The monitoring cost functions are easily understood. Since there is a critical level \hat{l}_L (possibly larger than 1 such as in case c) whereby F attracts the clean firm for $l_L \leq \hat{l}_L$ and the dirty firm for $l_L > \hat{l}_L$, monitoring costs up to \hat{l}_L are given by $(1 + s_d)l_L$ and above by l_L . As the figure illustrates, the sharing tax limit varies considerably with parameters k , s_d and l_L . In the following, we discuss how this variation comes about and how it leads to the four different types of outcome.

First, suppose that the *mcpe* is relatively low. To illustrate what happens if s_d increases, keeping k fixed, let us compare case (a) with case (b) in Figure 4. Suppose that L sets a large ES level to attract the clean firm. For large s_d (case b), that is if the dirty firm pollutes at a high level, F can extract a high tax from the dirty firm since the high ES level in L has a strong repellent effect on this firm. At the same time, attracting both firms is expensive for F since the clean firm suffers a large negative externality in the presence of the dirty firm and is tempted by the high ES level in L .²⁵ This leads to a large sharing tax limit for high levels of l_L . In contrast, for small s_d (case a), the dirty firm needs bigger tax incentives to stay in F because the high ES level in L is not very costly for it to abide by, and reduces the externality from the home firm. At the same time, for low k and s_d , there is no strong force that keeps the two mobile firms apart and attracting both firms is not very costly. Therefore the sharing tax limit is small for small s_d . In summary, if the *mcpe* is relatively small, a large ES level pays off for L if s_d is sufficiently large, but not otherwise.

²⁵This argument only holds if k is not too small, which is why there is the region to the left of the wedge-shaped region (b) in Figure 2.

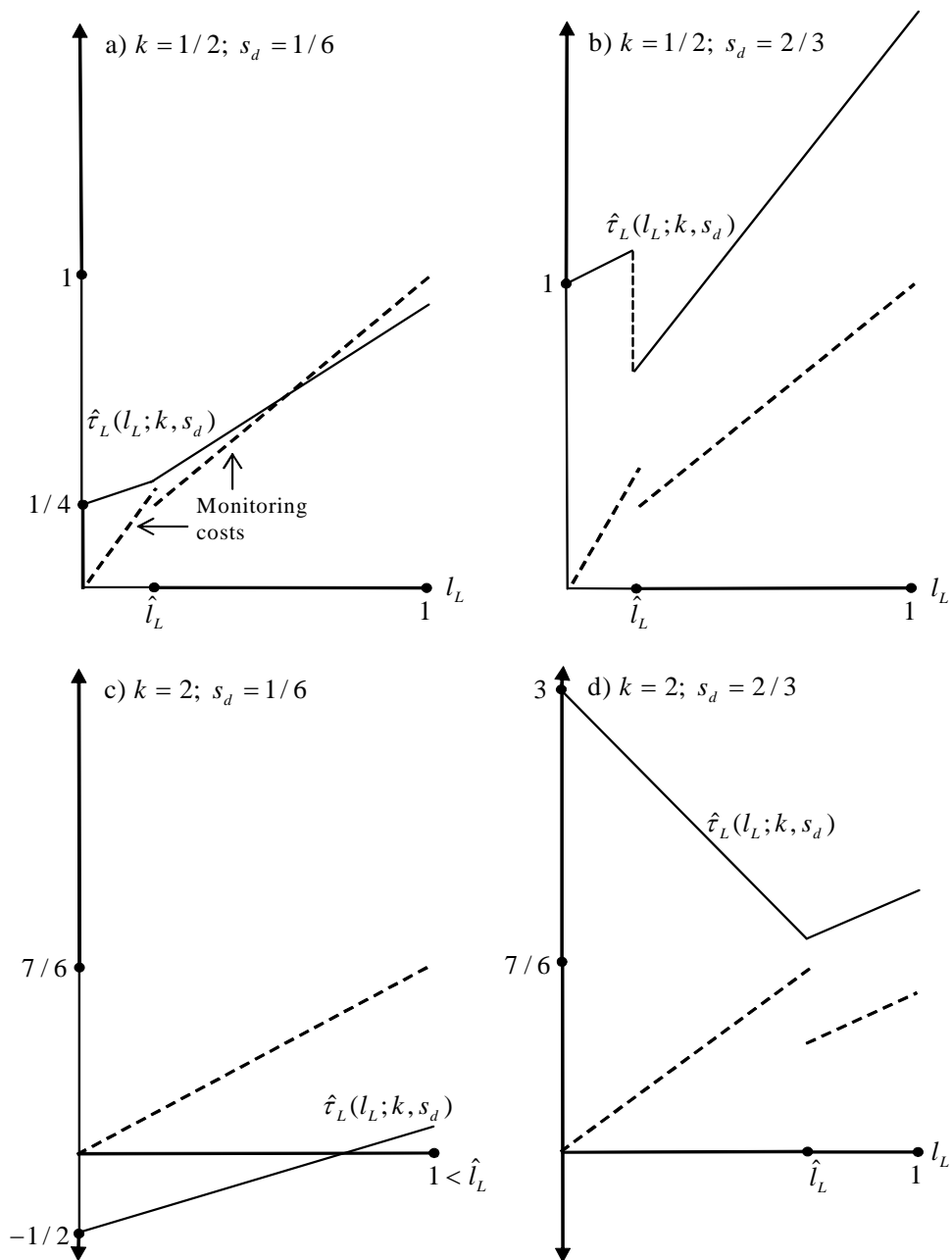


Figure 4

Now suppose that the *mcp*e is relatively high. The critical level of s_d above which it pays off for L to attract the clean firm (\hat{s}_d in Figure 2) depends on k as well. In particular, for sufficiently large k , this critical level is never attained (case d). If L sets a low ES to attract

the dirty firm, the large k and s_d together with this low ES level have a strong repellent effect on the clean firm so its willingness to pay to locate in F is large. At the same time, attracting both firms requires the highest ES level (because of the large $mcpe$), and a low tax, because the high s_d gives the dirty firm a strong incentive to locate in L where the ES level is low. Thus the sharing tax limit is high at low levels of l_L . In contrast, setting a large ES level to induce F to attract the dirty firm is difficult since for sufficiently large k even the dirty firm prefers, all else equal, larger ES levels. So F would set a high ES level as well and compete with L via taxes. However, if F already sets a high ES level and a competitive tax, it might just as well in addition attract the clean firm (and since the clean firm does not need to be monitored, this is not costlier than attracting only the dirty firm). Thus the sharing tax limit is low at high levels of l_L . Overall L is better off by setting a low ES.

Lastly, for low s_d and sufficiently large k (case c), L cannot make positive rents. This is because, for F , attracting the clean firm dominates attracting the dirty firm at all levels of l_L . However, attracting the clean firm (by setting the highest ES level) in turn is dominated by undercutting because, due to the low s_d , the government of F can attract the dirty firm as well without having to lower its tax substantially. As a consequence, L induces undercutting to avoid negative rents.

5 Relation to the Literature and Conclusions

We began this paper by noting concerns in policy circles that developing countries resembling those of recent entrants to the EU may, under certain circumstances, have a second-mover advantage in setting ESs and taxes. We then set out a formal framework which makes precise a set of circumstances under which such a second-mover advantage may arise. Four possible predictions were made about the outcome of fiscal competition when the public good in question is an ES. The particular prediction that emerges in equilibrium depends on the $mcpe$ and on the difference between mobile firms in their levels of pollution captured by s_d . The model focuses on the interplay between governments' incentives to manipulate policy - ESs and taxes - in order to maximize rents and firms' incentives to locate where these policies and the environmental externalities emitted by other firms have the most favorable impact on their profits.

If the $mcpe$ and s_d are both low, then environmental externalities and ESs are not important enough to firms for governments to be able to use them strategically. In this case, the

forces of tax competition envisaged by Brennan and Buchanan dominate, and the outcome is efficient. If $mcpe$ is low but s_d is relatively high then the possibility of a pollution haven arises where the dirty firm locates in the developing country in order to escape excessively high ESs set in the developed country. If $mcpe$ is in an intermediate range then the possibility of overprovision of ESs arises but in this region the roles of the countries are reversed; the developed country sets no ES at all, putting itself in the position of the pollution haven, while the developing country does set its ES at a positive level. For relatively high s_d in this range the developed country is successful in attracting one firm, the dirty firm, while for low s_d the developed country is unable to attract any firm. For $mcpe$ at a relatively high level, the equilibrium configurations do not change from the intermediate range but the implications for efficiency do; underprovision becomes a possibility, as does efficiency.

We will now place the paper's contribution to the literature, starting with the literature that follows Tiebout (1956) and then moving on to consider the literature that parallels fiscal federalism and standard tax competition. As in the literature that follows Tiebout, governments in our model are rent (or profit) maximizing but are constrained by competition. For example, Fischel (1975) and White (1975) share with the present paper the assumption that there is variation over firms' preferences for standards. In keeping with our model, there are no cross-border externalities. In contrast to our model, Fischel (1975) and White (1975) both assume that individual firms can be targeted for transfers and there is 'free entry' of jurisdictions, none of which has sufficient market power to extract rents from firms. As a result, within such a setting, an efficient outcome can be demonstrated in which firms 'vote with their feet.' In our model firms cannot be targeted for transfers. There is policy failure in the sense that once the policies are set they cannot be altered. And there are only two 'jurisdictions.' It is interesting to note that none of these differences in modeling approach matter for the achievement of efficiency providing that the $mcpe$ and s_d are sufficiently small. It is only when the $mcpe$ or s_d becomes sufficiently large for governments to compete for firms using ESs that a divergence from efficiency may arise. Given the focus on efficient outcomes in this first area of the literature and the fact that the focus tends to be on cooperative frameworks, there is limited scope in such a setting for exploring the type of second-mover advantage that is our focus in the present paper.

Let us turn now to the second area of the literature on competition over ESs and taxes,

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

In spite of the differences between the models in the second area and our model, the forces of competition between governments can operate in a similar way. This is seen most clearly by comparing the model of Markusen et al (1995) to ours. Recall that in Markusen et al (1995), two jurisdictions compete for the plants of a firm using two pollution taxes, one on domestic production and one on exports. The benchmark situation is one where the firm locates all its production facilities at home. Yet providing that transport costs are high enough and plant set-up costs are low enough, the foreign government can undercut the home government, much as in our model, to get some or all production to locate in its jurisdiction. The country that hosts the firm at the outset has an incentive to act in the manner of a limit-pricing monopolist, ‘limit-taxing’ the other country, by setting taxes just low enough that the firm is indifferent between locating its plants in one country or two and welfare is the same in both countries. (This effect is brought out most clearly by Levinson (1997), who rewrites the model of Markusen et al so that the monopoly rents are earned locally to where a plant locates.) Similarly, in our framework the developing country engages in limit-taxing by setting a tax just low enough to attract the firm(s) that it wishes to attract. In equilibrium, at least one firm is just indifferent between locating in the country that it chooses and moving to the other. Although it appears the basic underlying structure of the models in the second area of the literature could be adapted to study second-mover advantage, as far as we are aware this has not been done previously. The standard approach is to assume that both jurisdictions are ex ante symmetrical, externalities are internalized by the agents within each jurisdiction, and governments choose policies simultaneously without

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

communicating, yielding an outcome in which both governments obtain the same payoff in equilibrium.²⁷

It may be helpful to see the effects that motivate our results by direct comparison to the broader literature on fiscal competition. We are comparing the role of competition over taxes and public goods more conventionally defined against the role of competition over taxes and ESs in an environment where firms exert externalities on one another. The novel feature of competition in our model is that the developed country is compelled to set its ES and tax in such a way that it does not induce the developing country to undercut, and that this effect is strong enough to give the developing country a second-mover advantage. Our aim is to establish the novelty of this effect through comparison to the rest of the literature.

In a standard model of fiscal competition, Keen and Marchand (1997) consider a setting where the composition of public good provision matters. A jurisdiction can attract capital by increasing its productivity through a shift in spending from ‘public consumption goods’ such as parks and art galleries to ‘public investment goods’ such as road and communications networks. In this setting, competition in public good provision, which raises the return to capital that locates there, works in much the same way as tax competition. Thus, if the public good-capital complementarity is sufficiently strong, the equilibrium can exhibit over-provision of the public good in equilibrium. Although the setting of Keen and Marchand’s discussion invites a consideration of agglomeration externalities they leave these forces aside. Focusing on the forces of competition over public goods alone, there seems no reason to suppose that one jurisdiction or another should derive an advantage from this kind of competition, and ex ante symmetrical jurisdictions give rise to ex ante symmetrical outcomes for the respective jurisdictions in this setting.²⁸

Baldwin and Krugman (2004) do focus on agglomeration externalities. In their model,

²⁷Davies and Ellis (2007) do allow for the possibility that externalities between agents within a jurisdiction are not internalized. They show that in the presence of spillovers binding performance requirements, of which ESs are a special case, can act as a coordination device for firms. In equilibrium, jurisdictions choose performance requirements that maximize joint surplus from investment, which tax competition transfers to firms via tax subsidies. Thus equilibrium is efficient and governments do not make rents.

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

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

Our model shares similarities with Bucovetsky (2005) as well. He examines an interesting situation that blends key features of previous models; many jurisdictions may compete for mobile capital in a setting where agglomeration externalities are important. To focus on the special features of competition over public investment goods, or ‘public inputs’ as he refers to them, he abstracts from tax competition entirely. If agglomeration externalities are sufficiently strong, the efficient solution has all of the mobile factor locating in one jurisdiction. Yet competition among jurisdictions can lead to a Nash equilibrium in which the mobile factor locates across more than one jurisdiction, in contrast to Baldwin and Krugman and similar to our model. Competition between jurisdictions can lead to over-provision of the public good as in our model. A nice feature of the model which ours does not share is that comparative statics can be carried out whereby the economies of scale can be increased by an increase in the degree of substitutability between the goods that jurisdictions produce, potentially bringing about a decrease in the number of jurisdictions that provide public investment goods in equilibrium. Our model is different in the specific mechanism by which ESs are set either excessively high or excessively low. In Bucovetsky over-provision of the public input arises as jurisdictions fail to take into account the negative externalities that an increase in the public input inflicts on other jurisdictions. In our model

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

there are no such cross-border externalities. Instead, the developing country's second-mover advantage leads the developed country to set an inefficiently high or low ES to induce less aggressive competition by the developing country.

Inevitably, the theoretical framework developed here simplifies the situation in a number of key respects. The most noticeable simplification we have made is that we have not explicitly treated consumers in our analysis. As mentioned in the Introduction, most of the prior work on governments' strategic interactions over ESs has been motivated by a desire to satisfy consumers, giving rise to the NIMBY phenomenon. A possible direction for future research would be to integrate our approach with the approach of prior work to examine the interaction of the two competing forces. One potential limitation to our result is that in no outcome does the government of the developing country set ESs 'too low.' While it seems reasonable to argue that developed countries may set ESs too high, a concern is that developing countries actually set their ESs too low from the perspective of consumers. The introduction of consumers to the model could make it possible for ESs to be set too low in the developing country.

A potential limitation of our result is that the developing country always sets taxes higher than the developed country, which is counter-factual particularly to the core European countries versus the V4 which form our benchmark example (World Bank 1996). In the Introduction we argued that the developing country would set lower taxes if we changed the model in such a way that the mobile firms were initially located in the developed country and had an attachment to home. An alternative approach to reach a similar outcome in a more nuanced way would be to combine our framework with that of Keen and Marchand (1997), substituting their competition over public consumption goods with our competition over ESs. In the new framework the developed country would trade off competition for firms by making them more productive using 'public inputs' such as roads and communications networks but requiring them to pay higher taxes, as in Keen and Marchand, against a desire not to be undercut by the developing country. While higher taxation and public input provision could be observed in the developed country, a second-mover advantage could be preserved in the developing country because the latter would not undertake expenditures on public inputs.

Another direction for future research would be to ask how robust our results would be to the introduction of a larger number of countries and firms to the model. From our analysis

it is not obvious how the outcomes would be changed. One conjecture worth investigating 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.³⁰

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

³⁰It 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.

6 Appendix

Proof of Theorem 1.

Begin with the developing country, F , who has four basic strategies: A) attract only the clean firm, B) attract only the dirty firm, C) attract both the clean and the dirty firm, and D) attract neither of the mobile firms.

A) To attract only the clean firm, the following two conditions must hold

$$\begin{aligned}\tau_F + k(1 - l_F) &\leq \tau_L + k(1 - l_L)(1 + s_d) \quad \text{and} \\ \tau_L + s_d l_L + k(1 - l_L)(1 - s_d) &< \tau_F + s_d l_F + k(1 - l_F)(1 - s_d),\end{aligned}$$

from where we get that

$$\tau_F \in (\tau_L + (l_L - l_F)(s_d - k + ks_d), \tau_L + k(l_F + s_d - l_L(1 + s_d))].$$

For this interval to be nonempty we need $l_F > \frac{2kl_L + l_L - k}{1+k}$. Given a standard l_F , the follower would always set the highest possible tax. Thus the optimal standard level depends on the sign of

$$\begin{aligned}\frac{\partial(\tau_F - l_F)}{\partial l_F} &= \frac{\partial(\tau_L + k(l_F + s_d - l_L(1 + s_d)) - l_F)}{\partial l_F} \\ &= k - 1,\end{aligned}$$

and therefore, if $k \leq 1$, $l_F = \text{Max}\{0, \frac{2kl_L + l_L - k}{1+k}\}$ and if $k > 1$ then $l_F = 1$ (but for that we need $l_L < 1$). Note that at $k = 1$, the follower is indifferent across all ES levels. By assumption, he chooses the lowest one. This is the same for the following cases.

B) Similarly, to attract only the dirty firm, we need

$$\tau_F \in (\tau_L + k(l_F - l_L - s_d + l_F s_d), \tau_L + (l_L - l_F)(s_d - k + ks_d)].$$

For this interval to be nonempty we need $l_F < \frac{kl_L + l_L + k}{1+2k}$. The optimal standard level depends on the sign of

$$\begin{aligned}\frac{\partial(\tau_F - l_F)}{\partial l_F} &= \frac{\partial(\tau_L + (l_L - l_F)(s_d - k + ks_d) - (1 + s_d)l_F)}{\partial l_F} \\ &= -2s_d + k(1 - s_d) - 1,\end{aligned}$$

and therefore, if $k \leq \frac{1+2s_d}{1-s_d}$, then $l_F = 0$, and if $k > \frac{1+2s_d}{1-s_d}$, then $l_F = \frac{kl_L + l_L + k}{1+2k}$.

C) To attract both mobile firms, we need

$$\begin{aligned}\tau_F &\leq \text{Min} \{ \tau_F^1, \tau_F^2 \} \\ &\equiv \text{Min} \{ \tau_L + k(l_F - l_L - s_d + l_F s_d), \tau_L + (l_L - l_F)(s_d - k + k s_d) \}.\end{aligned}$$

We have $\tau_F^1 \leq \tau_F^2 \Leftrightarrow l_F \leq \frac{kl_L + l_L + k}{1+2k}$. Given a standard l_F , the optimal tax is $\tau_F = \text{Min} \{ \tau_F^1, \tau_F^2 \}$.

Case 1: $l_F \leq \frac{kl_L + l_L + k}{1+2k}$

$$\begin{aligned}\frac{\partial (2 * \tau_F^1 - (1 + s_d)l_F)}{\partial l_F} &= \frac{\partial (2\tau_L + 2k(l_F - l_L - s_d + l_F s_d) - (1 + s_d)l_F)}{\partial l_F} \\ &= 2k(1 + s_d) - (1 + s_d),\end{aligned}$$

and therefore, if $k \leq \frac{1}{2}$, then $l_F = 0$, and if $k > \frac{1}{2}$, then $l_F = \frac{kl_L + l_L + k}{1+2k}$.

Case 2: $l_F \geq \frac{kl_L + l_L + k}{1+2k}$

$$\begin{aligned}\frac{\partial (2 * \tau_F^2 - (1 + s_d)l_F)}{\partial l_F} &= \frac{\partial (2\tau_L + 2(l_L - l_F)(s_d - k + k s_d) - (1 + s_d)l_F)}{\partial l_F} \\ &= -3s_d - 1 + 2k(1 - s_d),\end{aligned}$$

and therefore, if $k \leq \frac{1+3s_d}{2(1-s_d)}$, then $l_F = \frac{kl_L + l_L + k}{1+2k}$, and if $k > \frac{1+3s_d}{2(1-s_d)}$, then $l_F = 1$. Putting the two cases together we obtain

$$\begin{aligned}l_F &= 0 \text{ and } \tau_F^1 \text{ if } k \leq \frac{1}{2}, \\ l_F &= \frac{kl_L + l_L + k}{1 + 2k} \text{ and } \tau_F^1 = \tau_F^2 \text{ if } k \in \left(\frac{1}{2}, \frac{1 + 3s_d}{2(1 - s_d)} \right], \text{ and} \\ l_F &= 1 \text{ and } \tau_F^2 \text{ if } k > \frac{1 + 3s_d}{2(1 - s_d)}.\end{aligned}$$

D) By setting a sufficiently high tax, F deters both mobile firms, earning zero rents. This never happens in equilibrium as the following argument shows. Note that rents in C are higher than in A or B for sufficiently large τ_L (the maximum taxes in A, B, and C increase by the same rate in τ_L and in C this tax is collected twice). Thus L can always induce C and make zero rents. Therefore, in equilibrium, L sets a tax that covers the costs of monitoring the standard. It is easy to show that for any τ_L sufficiently large to do that, F has at least one strategy (in A, B, or C) that makes positive rents, showing that F never chooses D.

The next step is to compare F 's rents across A, B, and C. According to the results so far, there are five different ranges of k to consider: If $s_d < \frac{1}{5}$, these are I) $k \leq \frac{1}{2}$, II)

$k \in \left(\frac{1}{2}, \frac{1+3s_d}{2(1-s_d)}\right]$, III) $k \in \left(\frac{1+3s_d}{2(1-s_d)}, 1\right]$, IV) $k \in \left(1, \frac{1+2s_d}{1-s_d}\right]$, and V) $k > \frac{1+2s_d}{1-s_d}$, and if $s_d \geq \frac{1}{5}$, these are I') $k \leq \frac{1}{2}$, II') $k \in \left(\frac{1}{2}, 1\right]$, III') $k \in \left(1, \frac{1+3s_d}{2(1-s_d)}\right]$, IV') $k \in \left(\frac{1+3s_d}{2(1-s_d)}, \frac{1+2s_d}{1-s_d}\right]$, and V') $k > \frac{1+2s_d}{1-s_d}$.

I and I') $k \leq \frac{1}{2}$. Rents in each case are.

$$\begin{aligned} r_A &= \begin{cases} \tau_L + k(s_d - l_L(1 + s_d)) & \text{if } l_L < \frac{k}{2k+1} \\ \tau_L + k\left(\frac{2kl_L + l_L - k}{1+k} + s_d - l_L(1 + s_d)\right) - \frac{2kl_L + l_L - k}{1+k} & \text{if } l_L \geq \frac{k}{2k+1} \end{cases}, \\ r_B &= \tau_L + l_L(s_d - k + ks_d), \text{ and} \\ r_C &= 2\tau_L - 2k(l_L + s_d). \end{aligned}$$

For $l_L \leq \frac{k}{2k+1}$, we have $r_A \geq r_B$ and for $l_L > \frac{k}{2k+1}$ we have $r_B > r_A$ (if $r_A = r_B$, F chooses A by assumption). Next, compare r_A with r_C for $l_L \leq \frac{k}{2k+1}$ and r_B with r_C for $l_L > \frac{k}{2k+1}$. This comparison yields

$$\begin{aligned} r_A \geq r_C &\Leftrightarrow \tau_L \leq k(l_L + 3s_d - l_L s_d), \text{ and} \\ r_B \geq r_C &\Leftrightarrow \tau_L \leq kl_L + 2ks_d + l_L s_d + kl_L s_d. \end{aligned}$$

Thus, for each case there is a sharing tax limit, an upper bound on the tax the leader can set without having the follower ‘undercut’ him. We denote this limit by $\hat{\tau}_L(l_L; k, s_d)$.

For the remaining ranges of k , the derivation is as in (I and I'), so we omit the details:

II and II') $k \in \left(\frac{1}{2}, \text{Min} \left\{ \frac{1+3s_d}{2(1-s_d)}, 1 \right\}\right]$. The sharing tax limit in this range of k is given by $\hat{\tau}_L(l_L; k, s_d) =$

$$\begin{cases} (2k+1)^{-1}(k + l_L + 4ks_d + l_L s_d - 2kl_L s_d - 2k^2 + 4k^2 s_d - 4k^2 l_L s_d) & \text{if } l_L \leq \frac{k}{2k+1} \\ (2k+1)^{-1}(k + l_L + 3ks_d + 2l_L s_d + 2kl_L s_d - 2k^2 + 2k^2 s_d) & \text{if } l_L > \frac{k}{2k+1} \end{cases}.$$

III) $k \in \left(\frac{1+3s_d}{2(1-s_d)}, 1\right]$. Here, $\hat{\tau}_L(l_L; k, s_d) =$

$$\begin{cases} 3s_d - 2k + kl_L + 3ks_d - 2l_L s_d - 3kl_L s_d + 1 & \text{if } l_L \leq \frac{k}{2k+1} \\ 3s_d - 2k + kl_L + 2ks_d - l_L s_d - kl_L s_d + 1 & \text{if } l_L > \frac{k}{2k+1} \end{cases}.$$

III') $k \in \left(1, \frac{1+3s_d}{2(1-s_d)}\right]$. Here, $\hat{\tau}_L(l_L; k, s_d) =$

$$\begin{cases} -(2k+1)^{-1}(2kl_L s_d - 4ks_d - l_L s_d - l_L - 4k^2 s_d + 4k^2 l_L s_d + 1) & \text{if } l_L \leq \frac{-1+k+ks_d}{s_d(1+2k)} \\ (2k+1)^{-1}(k + l_L + 3ks_d + 2l_L s_d + 2kl_L s_d - 2k^2 + 2k^2 s_d) & \text{if } l_L > \frac{-1+k+ks_d}{s_d(1+2k)} \end{cases}.$$

IV and IV') $k \in \left(\text{Max} \left\{ \frac{1+3s_d}{2(1-s_d)}, 1 \right\}, \frac{1+2s_d}{1-s_d} \right]$. Here, $\hat{\tau}_L(l_L; k, s_d) =$

$$\begin{cases} -(k - 3s_d - kl_L - 3ks_d + 2l_Ls_d + 3kl_Ls_d) & \text{if } l_L \leq \frac{-1+k+ks_d}{s_d(1+2k)} \\ 3s_d - 2k + kl_L + 2ks_d - l_Ls_d - kl_Ls_d + 1 & \text{if } l_L > \frac{-1+k+ks_d}{s_d(1+2k)} \end{cases}.$$

V) and V') $k > \frac{1+2s_d}{1-s_d}$. In this range of k , we have $r_A > r_B$ for all $l_L < 1$ and the sharing tax limit is given by $\hat{\tau}_L(l_L; k, s_d) =$

$$-(k - 3s_d - kl_L - 3ks_d + 2l_Ls_d + 3kl_Ls_d).$$

Here, if $l_L = 1$, then A is not feasible for F . Thus for $l_L = 1$ we compare r_B with r_C , which yields $\hat{\tau}_L(l_L; k, s_d) = 0$. We can ignore this case in the following since we have already established that L makes at least zero rents.

Now consider the developed country, L . Since F 's optimal response depends on the value of k , we need to consider L 's problem for each of the cases I-IV (I'-V'). To make positive rents L tries to induce F to choose a strategy in A or B. The maximum tax L can set at a given ES level is the corresponding sharing tax limit. It only remains to determine L 's optimal ES level. Except for V and V', the sharing tax limit has one discontinuity, where F switches from A to B. We derive the optimal ES level in three steps. First determine, what are the optimal ES levels to induce A and to induce B, second, compare rents at these levels, and third determine when L 's rents are negative. In this case, L induces C by setting $l_L = 0$ and some tax τ_L above the corresponding sharing tax limit.

I and I') $k \leq \frac{1}{2}$.

(i) $l_L \leq \frac{k}{2k+1}$. To determine the optimal ES level, consider the sign of

$$\begin{aligned} \frac{\partial (\hat{\tau}_L(l_L; k, s_d) - (1+s_d)l_L)}{\partial l_L} &= \frac{\partial (k(l_L + 3s_d - l_Ls_d) - (1+s_d)l_L)}{\partial l_L} \\ &= k - s_d - ks_d - 1. \end{aligned}$$

Thus, conditional on $l_L \leq \frac{k}{2k+1}$ and in the relevant range of k , the leader sets $l_L = 0$.

(ii) $l_L > \frac{k}{2k+1}$. Consider the sign of

$$\begin{aligned} \frac{\partial (\hat{\tau}_L(l_L; k, s_d) - l_L)}{\partial l_L} &= \frac{\partial (kl_L + 2ks_d + l_Ls_d + kl_Ls_d - l_L)}{\partial l_L} \\ &= k + s_d + ks_d - 1. \end{aligned}$$

Thus, conditional on $l_L > \frac{k}{2k+1}$ and $k \leq \frac{1-s_d}{1+s_d}$ the leader sets $l_L = \frac{k}{2k+1}$, and if $k > \frac{1-s_d}{1+s_d}$ the leader sets $l_L = 1$. (We ignore for the moment that $l_L = \frac{k}{2k+1}$ is not in the relevant interval for

l_L . This does not play a role because when comparing rents in (i) and (ii) $l_L = \frac{k}{2k+1}$ is never chosen even though, for any $\varepsilon > 0$, it yields slightly higher rents than setting $l_L = \frac{k}{2k+1} + \varepsilon$. If indifferent, L chooses the lower ES level by assumption.

Comparing rents in (i) and (ii) yields that for $k \leq 1 - s_d$, the leader induces A , and otherwise B . We obtain

$$\begin{aligned} \text{If } k &\leq 1 - s_d, \\ l_L^* &= 0, \tau_L^* = 3ks_d \text{ and } l_F^* = 0, \tau_F^* = 4ks_d. \\ \text{If } k &> 1 - s_d, \\ l_L^* &= 1, \tau_L^* = k + 3ks_d + s_d \text{ and } l_F^* = 0, \tau_F^* = 4ks_d + 2s_d. \end{aligned}$$

For the remaining ranges of k , the derivation is as in (I and I'), so we omit the details:

II and II') $k \in \left(\frac{1}{2}, \text{Min} \left\{ \frac{1+3s_d}{2(1-s_d)}, 1 \right\}\right]$. We obtain

$$\begin{aligned} \text{If } k &< \frac{1}{4}s_d^{-1} \left(s_d + \sqrt{16s_d^2 + (2 - s_d)^2 - 2} \right), \\ l_L^* &= 1, \tau_L^* = (2k + 1)^{-1} (k + 1 + 5ks_d + 2s_d - 2k^2 + 2k^2s_d), \text{ and} \\ l_F^* &= 0, \tau_F^* = (2k + 1)^{-1} (k + 1 + 5ks_d + 2s_d - 2k^2 + 2k^2s_d) + (s_d - k + ks_d). \\ \text{If } k &\geq \frac{1}{4}s_d^{-1} \left(s_d + \sqrt{16s_d^2 + (2 - s_d)^2 - 2} \right), \\ l_L^* &= 0, \tau_L^* = (2k + 1)^{-1} (k + 4ks_d - 2k^2 + 4k^2s_d), \text{ and} \\ l_F^* &= 0, \tau_F^* = (2k + 1)^{-1} (k + 4ks_d - 2k^2 + 4k^2s_d) + ks_d. \end{aligned}$$

III) $k \in \left(\frac{1+3s_d}{2(1-s_d)}, 1\right]$. We obtain

$$\begin{aligned} l_L^* &= 0, \tau_L^* = 3s_d - 2k + 3ks_d + 1, \text{ and} \\ l_F^* &= 0, \tau_F^* = 3s_d - 2k + 3ks_d + 1 + ks_d. \end{aligned}$$

III') $k \in \left(1, \frac{1+3s_d}{2(1-s_d)}\right]$. We obtain

$$\begin{aligned} l_L^* &= 0, \tau_L^* = (2k + 1)^{-1} (4ks_d + 4k^2s_d - 1), \text{ and} \\ l_F^* &= 1, \tau_F^* = (2k + 1)^{-1} (4ks_d + 4k^2s_d - 1) + k(1 + s_d). \end{aligned}$$

IV and IV') $k \in \left(\text{Max} \left\{ \frac{1+3s_d}{2(1-s_d)}, 1 \right\}, \frac{1+2s_d}{1-s_d}\right]$. We obtain

$$\begin{aligned} l_L^* &= 0, \tau_L^* = 3s_d + 3ks_d - k \\ l_F^* &= 1, \tau_F^* = 3s_d + 4ks_d. \end{aligned}$$

V) and V') $k > \frac{1+2s_d}{1-s_d}$

Consider the sign of

$$\frac{\partial (\hat{\tau}_L(l_L; k, s_d) - (1 + s_d)l_L)}{\partial l_L} = k - 3s_d - 3ks_d - 1.$$

Therefore if $s_d \geq \frac{1}{3}$ or if $k \leq \frac{1+3s_d}{1-3s_d}$ (only possible if $s_d < \frac{1}{3}$) the leader sets $l_L = 0$ and if $s_d < \frac{1}{3}$ and $k > \frac{1+3s_d}{1-3s_d}$ the leader sets $l_L = 1$ (however, for $l_L = 1$, A is not feasible for F , but this does not matter since we have already excluded this case). We obtain (for now and ignoring the problem at $l_L = 1$)

$$\begin{aligned} \text{If } s_d &\geq \frac{1}{3} \text{ or if } k \leq \frac{1+3s_d}{1-3s_d}, \\ l_L^* &= 0, \tau_L^* = 3s_d + 3ks_d - k \text{ and } l_F^* = 1, \tau_F^* = 3s_d + 4ks_d. \\ \text{If } s_d &< \frac{1}{3} \text{ and } k > \frac{1+3s_d}{1-3s_d}, \\ l_L^* &= 1, \tau_L^* = s_d \text{ and } l_F^* = 1, \tau_F^* = s_d. \end{aligned}$$

If L 's rents are negative, it is better off setting $l_L = 0$ and a 'detering tax' of $\tau_L^d > \hat{\tau}_L(0; k, s_d)$ to induce F to undercut, thereby ensuring zero rents. This happens if $k \leq 1$ and $s_d < \frac{2k-1}{3(k+1)}$ and if $k > 1$ and $s_d < \frac{k}{3(k+1)}$. In these cases F 's best response is to set $l_F = 1$ and $\tau_F = \tau_L^d - (s_d - k + ks_d)$.

Given these equilibrium strategies, it is straightforward to verify that there is always a second-mover advantage. \square

References

- [1] Andanova, L.B. (2003); *Transnational Politics of the Environment: The European Union and Environmental Policy in Central and Eastern Europe*. MIT Press, Cambridge, MA.
- [2] Antweiler, A., B. Copeland and M.S. Taylor (2001); "Is Free Trade Good for the Environment?" *American Economic Review*, 91(4): 877-908.
- [3] Baldwin, R. and P. Krugman (2004); "Agglomeration, Integration and Tax Harmonization." *European Economic Review*, 48: 1-23.
- [4] Black, D.A. and W.H. Hoyt (1989); "Bidding for Firms." *American Economic Review*, 79(5): 1249-1256.
- [5] Boadway, R., K. Cuff and N. Marceau (2004); "Agglomeration Effects and the Competition for Firms." *International Tax and Public Finance* 11: 623-646.
- [6] Brennan, G. and J.M. Buchanan (1980); *The Power To Tax: Analytical Foundations of a Fiscal Constitution*, Cambridge University Press, Cambridge.
- [7] Brueckner, J.K. (2000); "A Tiebout/Tax Competition Model." *Journal of Public Economics*, 77: 285-306.
- [8] Bucovetsky, S. (2005); "Public Input Competition." *Journal of Public Economics*, 89: 1763 - 1788
- [9] Cumberland, J.H. (1979); "Interregional Pollution Spillovers and Consistency of Environmental Policy." Published in H. Siebert et al. eds., *Regional Environmental Policy: The Economic Issues*, New York University Press, New York.
- [10] Cumberland, J.H. (1981); "Efficiency and Equity in Interregional Environmental Management." *Review of Regional Studies* 2: 1-9.
- [11] Davies, R.B. and C.J. Ellis (2007); Competition in Taxes and Performance Requirements for Foreign Direct Investment." *European Economic Review*, 51(5): 1303-1325.

- [12] Ederington, J., A. Levinson and J. Minier (2004); "Trade Liberalization and Pollution Havens." Published in *Advances in Economic Analysis and Policy*, Vol. 4, No. 2, Article 6, the Berkley Electronic Press.
- [13] Euractiv (2010); "Commission proposes phasing out coal subsidies." <http://www.euractiv.com/en/energy/commission-proposes-phasing-out-coal-subsidies-news-496532>, published on July 21, 2010, last accessed on September 29, 2010.
- [14] 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.
- [15] Graham, E.M. (2000); *Fighting the Wrong Enemy: Antiglobal Activists and Multinational Enterprises*. Institute for International Economics, Washington D.C.
- [16] Groenert, V., M. Wooders and B. Zissimos (2009); "Developing Country Second-Mover Advantage in Competition over Standards and Taxes." Vanderbilt University Department of Economics Working Paper No. 09-W09.
- [17] Haufler, A. and I. Wooton (1999); "Country Size and Tax Competition for Foreign Direct Investment." *Journal of Public Economics*, 71: 121-139.
- [18] Justman, M., J.-F. Thisse, and T. van Ypersele (2002); "Taking the Bite out of Fiscal Competition." *Journal of Urban Economics* 52, 294–315.
- [19] Keen, M. and M. Marchand (1997); "Fiscal Competition and the Pattern of Public Spending." *Journal of Public Economics* 66: 33-53.
- [20] Kempf, H. and S. Rossignol (2007); "Is Inequality Harmful For The Environment In A Growing Economy?" *Economics and Politics*, 19(1): 53-71.
- [21] Kind, H., K. Knarvik and G. Schjelderup (2000); "Competing for Capital in a 'Lumpy' World." *Journal of Public Economics*, 78: 253-274.
- [22] Levinson, A. (1997); "A Note on Environmental Federalism: Interpreting Some Contradictory Results." *Journal of Environmental Economics and Management* 33: 359-366.
- [23] Levinson, A. (1999); "State Taxes and Interstate Hazardous Waste Shipments." *American Economic Review*, 89(3): 666-677.

- [24] Levinson, A. (2003); "Environmental Regulatory Competition: A Status Report and Some New Evidence." *National Tax Journal*, 56(1): 91-106.
- [25] 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.
- [26] New York Times (2010); "Europe's Enduring Coal Subsidies." <http://green.blogs.nytimes.com/2010/06/28/europes-enduring-coal-subsidies/>, article by J. Kanter, published on June 28, 2010, last accessed on October 7, 2010.
- [27] Oates, W.E. (1972); *Fiscal Federalism*, Harcourt Brace Jovanovich, New York.
- [28] 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.
- [29] Taylor, M.S. (2004); "Unbundling the Pollution Haven Hypothesis." Published in *Advances in Economic Analysis and Policy*, the Berkley Electronic Press.
- [30] Tiebout, J.M. (1956); "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64, 416-424.
- [31] 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.
- [32] Wilson, J.D. (1986); "A Theory of Interregional Tax Competition" *Journal of Urban Economics*, 19: 296-315.
- [33] 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.
- [34] Zodrow, G.R. and P. Mieszkowski (1986); "Pigou, Tiebout, Property Taxation and the Under-Provision of Local Public Goods." *Journal of Urban Economics*, 19: 356-370.