

Bank of Finland Research Discussion Papers
15 • 2023

Topi Hokkanen

Optimal carbon leakage



Bank of Finland
Research

Bank of Finland Research Discussion Papers
Editor-in-Chief Esa Jokivuolle

Bank of Finland Research Discussion Papers 15/2023
25 October 2023

Topi Hokkanen:
Optimal carbon leakage

ISBN 978-952-323-449-9, online
ISSN 1456-6184, online

Bank of Finland
Research Unit

PO Box 160
FIN-00101 Helsinki

Phone: +358 9 1831

Email: research@bof.fi

Website: www.suomenpankki.fi/en/research/research-unit/

The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Bank of Finland.

Optimal carbon leakage

Topi Hokkanen*

October 23, 2023

Abstract

Carbon leakage is one of the major issues facing policymakers today when designing environmental regulation. While the empirical and trade literature on carbon leakage is rich, much less is known about the implications of carbon leakage risk on optimal regulatory policies under asymmetric information. To this end, I derive the optimal incentive compatible mechanism to regulate polluting firms under asymmetric information of both their abatement costs and carbon leakage risk, which I model as type-dependent outside options. The resulting regulatory distortions depend on the affiliation between the firm's abatement and relocation costs. The optimal policy is less strict than first-best whenever this affiliation is negative or mildly positive, whereas under strong positive affiliation I find a novel upwards distortion in the optimal policy. My results imply that rather than being a byproduct of unsuccessful regulation, carbon leakage may be the optimally induced outcome of incentive compatible regulation, contrasting with the received wisdom in policy debate.

JEL classification: D62, D82, L51, Q54, Q58.

Keywords: carbon leakage, mechanism design, externalities, asymmetric information.

*Bank of Finland, Payment Systems Department and Aalto School of Business, Department of Economics. This paper has benefited greatly from conversations with Matti Liski, Juuso Välimäki, and Andrey Zhukov. I am also grateful to Esa Jokivuolle, Jan Jääskeläinen, Deniz Kattwinkel, Klaus Kultti, Torben Mideksa, Juan-Pablo Montero, Teemu Pekkarinen, and Tuomas Takalo for helpful comments on earlier versions of this paper. The opinions expressed in this paper represent the author's own views and not necessarily those of the Bank of Finland.

1 Introduction

A common argument heard from firms and politicians alike when new regulation (such as environmental regulation, higher corporate tax rates, etc.) is considered is that the new regulation will only serve to drive domestic firms to relocate to other countries or areas with less stringent regulation, harming the domestic economy. With environmental externalities such as carbon dioxide, this threat is called *carbon leakage* and the concern for leakage is often referred to as the *pollution haven hypothesis* (see Levinson and Taylor (2008)). The European Commission (EC) defines carbon leakage as: “[a] situation that may occur if, for reasons of costs related to climate policies, businesses were to transfer production to other countries with laxer emission constraints. This could lead to an increase in their total emissions”¹. In practice, this loss of competitiveness and risk of carbon leakage is taken seriously enough that many regulatory mechanisms, such as the EU ETS make special exceptions for “especially vulnerable industries”, which are placed on a *carbon leakage list*. The exceptions made for these industries can include, e.g. direct subsidies, freely allocated emission permits or even full exemption from regulation. In order to fight leakage, EU recently implemented the Carbon Border Adjustment Mechanism (CBAM), which aims to minimize carbon leakage to lesser regulated countries, addressing it by imposing border carbon adjustments for imported goods².

In this paper, I show that in a simple mechanism design problem with firm-specific externalities, the key question isn’t simply *if firms relocate* due to stricter regulation, it’s *who relocates* and that carbon leakage, rather than being a sign of failed environmental policy may actually be *optimally induced* by the regulatory mechanism to save on socially costly public funds. I find a unique upward distortion in the abatement threshold (or, alternatively, carbon price) alongside the more commonly found downward distortions due to information rents and incentive compatibility. The key parameter driving the size and direction of these distortions is how the firms’ abatement costs associate with their relocation propensities, which remains an open empirical question³. Due to the ambiguity of the received empirical literature with regards to the correlation between these two cost parameters, I take no stance in this paper as to the most likely sign of this correlation, but instead set up an agnostic model that allows for any sign

¹<https://climate.ec.europa.eu/eu-action/eu-emissions-trading-system-eu-ets/free-allocation/carbon-leakage>

²<https://taxation-customs.ec.europa.eu/carbon-border-adjustment-mechanism>.

³See e.g. Ederington, Levinson, and Minier (2005).

and size of dependence or association between the two.

Formally, the model used in this paper makes use of *type-dependent outside options* in an otherwise relatively standard mechanism design problem with externalities. A previous application of type-dependent outside options interpreted as firm relocation costs in an asymmetric information regulation model appeared in Vislie (2000), but unlike in this paper he only considers the case where the *relocation costs* of firms are negatively correlated with their abatement costs⁴. In the carbon leakage model considered in this paper, I find similar results for the regulatory distortions in the abatement threshold under different association regimes, although countervailing incentives do not arise in my model as they do in Lewis and Sappington (1989) (or as defined by Aguirre and Beitia (2017)), even with type-dependent outside options.

While environmental regulation is a topic often covered in mechanism design (see, e.g. Montero (2008), Kim and Chang (1993), Dasgupta, Hammond, and Maskin (1980), Baliga and Maskin (2003) or Lewis and Sappington (1995)), carbon leakage and especially the mechanism design approach to the problem is novel, with one of the only preceding papers being Ahlvik and Liski (2021). The model in this paper shares essentially the same set-up as Ahlvik and Liski (2021), who consider the firm relocation (carbon leakage) issue in a mechanism design framework with global externality-producing firms, but in a more general setup without explicitly specifying the mapping between relocation costs and abatement costs. By applying the random participation mechanism of Rochet and Stole (2002) to make the model tractable, they arrive at the same novel upwards distortion in the abatement threshold as I do in my more stylized model. This paper benefits greatly from a similar setup, but allows for cleaner results due to the additional - although more restrictive - assumption of a one-to-one mapping between the abatement costs and the relocation costs⁵.

The majority of empirical literature on carbon leakage has focused on either estimating carbon leakage rates within specific industries under unilateral regulation (such as the EU-ETS), often assessing its effects on (different measures of) industry competitiveness. While this literature in its entirety is much too vast for me to parse in this

⁴This is a very similar setup as the $k < 0$ case in this paper.

⁵Without this mapping, the problem would essentially be a multidimensional screening problem, which are notoriously difficult to solve generally. A brief survey of these problems can be found in, e.g. Armstrong and Rochet (1999).

introduction, I will highlight some recent papers that are relevant to the model used in this paper. For a comprehensive survey of the empirical literature related to the competitiveness effects of the EU-ETS, I refer the reader to Verde (2020) or Dechezleprêtre and Sato (2017) or Venmans, Ellis, and Nachtigall (2020). To the best of my knowledge, there have been very few empirical papers that have tried to estimate the association between the abatement (or compliance) costs and relocation costs (or, alternatively, relocation propensities) of firms under regulation, which is a key parameter used in this paper⁶. One such paper is Ederington, Levinson, and Minier (2005), who find evidence of a positive association between abatement and relocation costs, whereas Levinson and Taylor (2008) find that industries whose abatement costs increased the most increased their net imports the most. Focusing on trade flows, Naegele and Zaklan (2019) find no evidence of EU ETS causing carbon leakage in European manufacturing whereas recent estimates by Misch and Wingender (2021) generally show high leakage rates across the European Union caused by the EU ETS. In contrast to the other leakage literature, Eskeland and Harrison (2003) test the pollution haven hypothesis and find that firms relocating to lesser regulated countries tended to be cleaner than their counterparts, polluting less.

⁶Ahlvik, Liski, and Martin (2017) do exactly this, using the survey data from Martin, Muûls, De Preux, and Wagner (2014)

2 The model

Consider a continuum of firms with unit mass, each characterized by a privately known abatement cost, $c \in [\underline{c}, \bar{c}]$ (with $\underline{c} \geq 0$) of reducing one unit of emissions. A single unit of emissions causes a global externality of size $D > 0$ that firms impose on the country where they reside, absent any abatement. Aside from negative externalities, each firm generates $\gamma > 0$ of social welfare with their activities, hence the need for the regulator to balance the gains and losses of regulating the firms. I assume that the abatement costs c are distributed according to a continuously differentiable density function $f(c)$, strictly positive on (\underline{c}, \bar{c}) , with cumulative distribution function $F(c)$ satisfying the monotone hazard rate assumption. The shadow cost of public funds is positive and equal to λ .

Assumption 1. *Monotone Hazard Rate (regularity of F): I assume that the distribution of types $F(\cdot)$ satisfies the monotone hazard rate assumption, i.e.*

$$\frac{d}{dc} \left(\frac{f(c)}{1 - F(c)} \right) \geq 0 \quad (1)$$

The initial location for each firm is the home country, i.e. country i . Each firm also has the opportunity to relocate to another country, the cost of which is given by a linear relocation cost

$$\theta := \bar{\theta} + kc \quad (2)$$

where $k, \bar{\theta} \in \mathbb{R}$, serving as a type-dependent outside option for the firm. The key parameter in this linear relationship is the association parameter $k \in \mathbb{R}$. If $k < 0$, then the abatement costs c and relocation costs θ are negatively associated, meaning that the highest cost abaters (those with the highest c) face the smallest relocation costs, or are the most mobile. Under positive association, i.e. when $k > 1$, we have the opposite, so that the lowest cost abaters are the most mobile, and the highest cost firms are the most immobile. In the intermediate case, when $0 < k < 1$, relocation costs are weakly positively associated with the relocation costs, with the lowest cost firms being the most mobile.

I denote the alternative location by j and for the most part, I assume that location j has no regulatory policies in place. The home country can retain firms by means of offering

them a mechanism denoted as $\mathcal{M} := \{X(c), T(c)\}$, consisting of:

Abatement $X(c) \in [0, 1]$, ie. whether a c -type firm cuts emissions or not, and

The transfer (subsidy) to the firm $T(c) \in \mathbb{R}$, given to the firm conditional on it staying in the home country.

Moreover, I require that the mechanism should be incentive compatible for the firms, meaning the firms should report their cost types truthfully. Firms are cost-minimizers, and choose to stay (i.e. report their cost type to the domestic mechanism) or relocate based on which action minimizes their compliance costs. Every firm has an option to relocate to the unregulated country, but this incurs a *relocation cost* θ , in (2).

The regulator wants to maximize the social surplus, ie. the surplus less firm compliance. The tools at its disposal are *incentive-compatible mechanisms* \mathcal{M} . The regulator has to take into account the type-dependent relocation cost θ , which later on gives us three cases to consider as far as distortions go.

2.1 The social welfare function and some preliminaries

I define the compliance cost of a firm in a given mechanism \mathcal{M} , as a linear function of their own private type, the abatement X and the subsidy T . When reporting a type c' , the compliance cost $C(c, c')$ of a firm of (privately known) type c is defined as

$$C(c, c') := cX(c') - T(c') \quad (3)$$

The firms' cost-minimization implies that the mechanism on offer should satisfy an incentive-compatibility constraint of the following type: reporting truthfully must minimize the compliance cost given the mechanism. Therefore we have the following IC constraint:

$$c \in \arg \min_{c' \in [c, \bar{c}]} C(c, c') \quad (4)$$

This constraint implies that in any incentive-compatible mechanism, the abatement $X(c)$ is a nonincreasing function of c .

Lemma 1. *An allocation $X(c)$ is incentive compatible iff it is a nonincreasing function.*

Proof. Take two types in $[\underline{c}, \bar{c}]$, say c and c' . Without loss of generality, let $c' > c$. Now the incentive compatibility constraints directly imply

$$cX(c) - T(c) \leq cX(c') - T(c') \quad (5)$$

$$c'X(c') - T(c') \leq c'X(c) - T(c) \quad (6)$$

Subtracting the constraints from one another and noting that we defined $c' > c$ yields $X(c') \leq X(c)$, which completes the proof. ■

The baseline for our model is that firms do not abate and stay in the country. Doing so, they impose a full negative externality of size D for the home country. If a firm abates their externality, it yields the home country a *climate surplus* of $(D - c)$. In addition to the possible climate surplus, the firm generates γ if it stays. Noting the shadow cost of public funds (λ), we can now define the social welfare function to be maximized by choice of mechanism, comprising of the benefit of retaining the firm less the firm compliance costs.

$$\begin{aligned} W(\mathcal{M}) &:= \int_{\underline{c}}^{\bar{c}} [\gamma + DX(c) - (1 + \lambda)T(c) - cX(c) + T(c)] f(c) dc \\ &= \int_{\underline{c}}^{\bar{c}} [\gamma + (D - c)X(c) - \lambda T(c)] f(c) dc \end{aligned} \quad (7)$$

The integrand is linear in the allocation $X(c)$, which moreover belongs to the interval $[0, 1]$. Therefore the optimal $X(c)$ must be a bang-bang solution, i.e. $X(c) \in \{0, 1\}$ for all c . Together with Lemma (1), which states that incentive compatibility requires $X(c)$ to be decreasing in c , we get that the incentive compatible allocation must be a step function, specifying a threshold type $c^* \in [\underline{c}, \bar{c}]$, such that every firm below or at the threshold is allocated full abatement, i.e. $X(c) = 1$, and firms above the threshold are allocated zero abatement:

$$X(c) = \mathbf{1}_{\{c \leq c^*\}} \quad (8)$$

In order to satisfy incentive compatibility, whenever the allocation $X(c)$ is constant, the transfer that corresponds with that allocation must also be a constant. Otherwise there would exist firm types that would find it profitable to misreport their types, altering their compliance costs, while keeping the allocation the same. Since the requirement for truthfully reporting c to be minimizer of $C(c, \hat{c})$, this immediately implies that the transfers must be constant whenever the allocation is a constant as well.

Lemma 2. *Whenever the allocation $X(c)$ is constant in an incentive compatible mechanism, the transfer $T(c)$ is constant.*

Proof. Suppose not. In this case there exist types $c, c' \in [\underline{c}, \bar{c}]$ for which

$$cX(c') - T(c') < cX(c) - T(c)$$

which contradicts incentive compatibility. ■

Therefore the optimal incentive compatible mechanism is fully characterized by a two part tariff with allocation $X(c)$ as defined above, and a transfer of $T(c) := T^* + \mathbf{1}_{\{c \leq c_i^*\}} c^*$. In other words, every firm gets a base transfer of T^* , and a top-up of c^* if they abate. Since I want to make the connection this stylized model has with Pigou taxation very clear, I rescale the mechanism \mathcal{M} by altering the transfers from

$$T(c) = T^* + \mathbf{1}_{\{c \leq c_i^*\}} c^* \tag{9}$$

to a downscaled transfer of

$$\hat{T}(c) := T^* - \left(1 - \mathbf{1}_{\{c \leq c_i^*\}}\right) c^* \tag{10}$$

and with the induced firm compliance cost of

$$\hat{C}(c) := \underbrace{c^*}_{\text{Carbon tax}} - \underbrace{T^*}_{\text{Base subsidy}} + \underbrace{X(c)(c - c^*)}_{\text{Incentive part}} \tag{11}$$

Lemma 3. *The rescaled mechanism is incentive-compatible.*

Proof. Firm compliance with the transfer above is

$$\hat{C}(c) = c^* - T^* + X(c)(c - c^*)$$

With the allocation being a step function, the IC constraints reduce to preventing deviations that cross the threshold c^* :

$$c - T^* \leq c^* - T^* \tag{12}$$

$$c^* - T^* \leq \hat{c} - T^* \tag{13}$$

for $c \leq c^* < \hat{c}$. These hold trivially for every T^* so the new mechanism is incentive-compatible. ■

This has the direct market interpretation so that c^* is both the threshold type, and *the price of carbon*. To see why, let us take a look at the compliance costs for abating and non-abating firms. Abating firms, for which $X(c) = 1$ have a compliance cost of $c - T^*$. That is, they abate fully, pay their private cost of that abatement and get a T^* base subsidy for this action. Non-abating firms face a compliance cost of $c^* - T^*$, likewise receiving the base subsidy of T^* , but now pay up front the *carbon tax* c^* due to their one unit of pollution (or zero units of abatement, respectively). Next, I take a look at the *relocation constraint* and formulate the optimization problem of the regulator.

2.2 The relocation constraint and regulator's problem

The regulator's objective function, given the two-part tariff implied by incentive compatibility and the rescaled transfer can now be written as

$$\max_{c^*, T^*} \underbrace{\int_c^{c^*} \{\gamma + (D - c)X(c) - \lambda T^*\} f(c) dc}_{\text{Welfare from abating firms}} + \underbrace{\int_{c^*}^{\bar{c}} (\gamma + \lambda(c^* - T^*)) f(c) dc}_{\text{Welfare from non-abating firms}} \tag{14}$$

subject to the individual rationality, or *relocation constraint* induced by (2). Non-abating

firms yield an additional⁷ λc^* of welfare, on top of γ , since they are taxed by the regulator for their emissions with a carbon price of c^* . Firms choose to relocate to the alternative location whenever their cost of compliance in the home country exceeds their relocation cost θ . That is, firms leave when

$$\hat{C}(c) > \bar{\theta} + kc \quad (15)$$

Therefore, the optimization problem of the regulator is to jointly decide $\{c^*, T^*\}$, given parameters $(D, \gamma, k, \bar{\theta})$ and the distribution of costs $F(c)$. However, the relocation constraint above induces a partition on the type space $[\underline{c}, \bar{c}]$, where some measure of firms will stay and abate, some measure will stay but not abate, and some measure leaves. The crucial parameter that induces this partition is the dependence parameter k .

2.3 The first-best benchmark

As a starting-off point, I will solve the regulator's problem when the regulator has perfect information, i.e. can simply optimize type-by-type. This will give us the efficient, or *first-best* outcome, to which I will later on compare our optimal mechanisms to.

In this case, the regulator can disregard the incentive compatibility constraints, and reduce every type c to its outside option payoff - in this case given by the relocation constraint in (2). Since incentive compatibility can be disregarded due to the firm's type being observable, the optimization problem of the regulator is to choose the allocation $X(c)$ to maximize

$$\max_{X(c)} [\gamma + DX(c) - (1 + \lambda)T(c) - \hat{C}(c)] \quad (16)$$

such that $\hat{C}(c) = \bar{\theta} + kc$ for all c . After inserting the compliance cost and simplifying, we have

⁷That is, I consider the shadow cost of public funds to be symmetric wrt. transfers. A tax from the firm to the regulator of c^* is worth $(1 + \lambda)c^*$ to the regulator

$$\max_{X(c)} [\gamma + (D - (1 + \lambda)c) X(c) + \lambda(\bar{\theta} + kc)] \quad (17)$$

where $X(c)$ is set to $X(c) = 1$ if $(D - (1 + \lambda)c) \geq 0$, yielding *the first-best* abatement threshold which I denote:

$$c^{FB} = \frac{D}{1 + \lambda}. \quad (18)$$

So we see that the abatement threshold is set at the socially optimal level, given that the regulator has a positive shadow cost for public funds⁸. The entire transfer schedule then is set such that it renders every type, regardless of abatement, at their outside option payoff. The transfer that implements this is

$$T(c) := \mathbf{1}_{\{c \leq c^{FB}\}} c - (\bar{\theta} + kc) \quad (19)$$

Lemma 4. *The perfect information mechanism is*

$$\begin{aligned} X(c) &= \mathbf{1}_{\{c \leq \frac{D}{1+\lambda}\}} \\ T(c) &= \mathbf{1}_{\{c \leq c^{FB}\}} c - (\bar{\theta} + kc) \end{aligned}$$

Proof. Noting that the only constraint relevant for the regulator's problem is the relocation constraint, simple inspection yields that the induced compliance cost for a firm of type c in the above mechanism is indeed $\hat{C}(c) = \bar{\theta} + kc$. ■

The perfect information mechanism retains all of the types, or alternatively induces full participation and hence implies no carbon leakage. This contrasts with the asymmetric information mechanisms soon to be derived, which generally induce limited participation, i.e. carbon leakage⁹.

⁸If the transfers to the firms would be costless, i.e. $\lambda = 0$, then we see that the threshold would be set at the full level of the externality, D . As we see later, this would also negate the asymmetric information problem as well.

⁹Vislie (2000), Laffont and Tirole (1993)

3 Local regulatory policies

I now derive the optimal policy for the regulator when they are regulating a global externality but have to tackle this issue by offering local regulatory policies (i.e. when the regulator cannot offer cross-border transfers to relocating firms) and moreover, when there is informational asymmetry between the firms and the regulator.

When the regulator cannot observe firm types, and can only offer transfers conditional on firms locating in the country, the mechanism design problem amounts to solving an optimization problem where the objective function is (14), subject to a specific partition of the type space, given by the relocation constraints in (15). I write out the relocation constraints explicitly below, and split the range of the association parameter k into three regimes, solving each separately. The relocation constraints with the induced compliance cost $\hat{C}(c)$ are, for a given c^* :

Assuming that firms above the threshold (non-abaters) relocate, i.e. **if $c > c^*$ leave**, then the partition is determined by:

$$c^* - T^* \geq \bar{\theta} + kc \quad (20)$$

$$-kc \geq \bar{\theta} + T^* - c^* \quad (21)$$

from where we get as the cut-off type

$$c \leq -\left(\frac{\bar{\theta} + T^* - c^*}{k}\right), \quad \text{for } k > 0 \quad (22)$$

$$c \geq -\left(\frac{\bar{\theta} + T^* - c^*}{k}\right), \quad \text{for } k < 0 \quad (23)$$

However, **if $c < c^*$ leave**, then the relocation constraint is

$$c - T^* \geq \bar{\theta} + kc$$

$$(1 - k)c \geq \bar{\theta} + T^*$$

giving us cut-off types of

$$c \leq \frac{\bar{\theta} + \Gamma^*}{1 - k}, \quad \text{for } k > 1 \quad (24)$$

$$c \geq \frac{\bar{\theta} + \Gamma^*}{1 - k}, \quad \text{for } k < 1 \quad (25)$$

For convenience, let us denote

$$c' := - \left(\frac{\bar{\theta} + \Gamma^* - c^*}{k} \right) \quad (26)$$

$$c'' := \frac{\bar{\theta} + \Gamma^*}{1 - k}. \quad (27)$$

These relocation constraints now induce a partition of staying and relocating types, depending on k . For clarity, I will analyze each regime for the association parameter k as a separate subsection.

3.1 $k < 0$: Negative association, where the low-cost firms are the most immobile

When $k < 0$, then we have a situation in which the firms with the highest relocation costs are ones with low abatement costs. We assume that $c' < \bar{c}$, so that we have an interior solution for the optimization problem¹⁰, and conjecture that the partition in this case is the following: firms with $c \leq c^*$ stay and abate, firms with $c^* < c \leq c'$ stay, but don't abate, and firms with $c > c'$ leave. The only relocation constraint that is relevant with this partition is (23), pinning down the marginal relocating type of firm above the abatement threshold c^* . We can now write the the social welfare function to be optimized as:

$$\max_{c^*, \Gamma^*} W = \int_{\underline{c}}^{c^*} (\gamma + (D - c) - \lambda \Gamma^*) f(c) dc + \int_{c^*}^{c'} (\gamma + \lambda(c^* - T^*)) f(c) dc \quad (28)$$

The optimal mechanism in this case is summarized in the following proposition:

¹⁰The corner solution is relegated to the appendix for readability.

Proposition 1. *Whenever $k < 0$, the low-cost firms are the most immobile in the home country, and the abatement threshold in the mechanism under global externalities is distorted lower than first-best, to*

$$c^* = \frac{D}{1 + \lambda} - \frac{\lambda}{1 + \lambda} \frac{F(c^*)}{f(c^*)}$$

Therefore I have a result of the threshold being distorted downwards from the first-best (c^{FB}), meaning the country sets a less strict regime than would be socially optimal. I illustrate the result below in Figure 1.

At the outset, one might not expect this to happen, since in this association regime the best firm types for the regulator (low-cost types) are the most immobile, and hence it might seem that the optimal mechanism should exploit this in favor of abatement. However, the downward distortion arises due to the dual role of c^* alluded to earlier, as in this stylized model it functions as both the abatement threshold *and* the carbon price, determining simultaneously both the efficient abatement cut-off *and* the price non-abating firms have to pay in order to remain in the home country. So, while this association regime has the low-cost firms face the highest relocation costs (since $k < 0$), these costs cannot be fully exploited¹¹ in the optimal mechanism (by increasing c^* to the socially optimal level), since this action would at the same time increase the compliance costs for all non-abaters (holding T^* constant), leading to more relocation (and loss of γ for that measure of firms).

Therefore, this is a restatement of a very classic result in mechanism design, namely the efficiency-rent extraction tradeoff. Similar results in incentive regulation can be found in Baron (1985), Laffont and Tirole (1996) or Laffont and Tirole (1993), where the optimal regulation falls short of the first-best level due to the requirement of maintaining incentive compatibility. Indeed, the result of downwards distortion in abatement arises also in Vislie (2000), where net abatement is below first-best levels (or, as modelled, the net emissions are distorted upwards) in order to reduce information rents flowing to the most efficient abaters, who otherwise all would have incentives to report a higher type.

¹¹In essence, the low-cost firms are the ones that are the most captive in the home country at the outset.

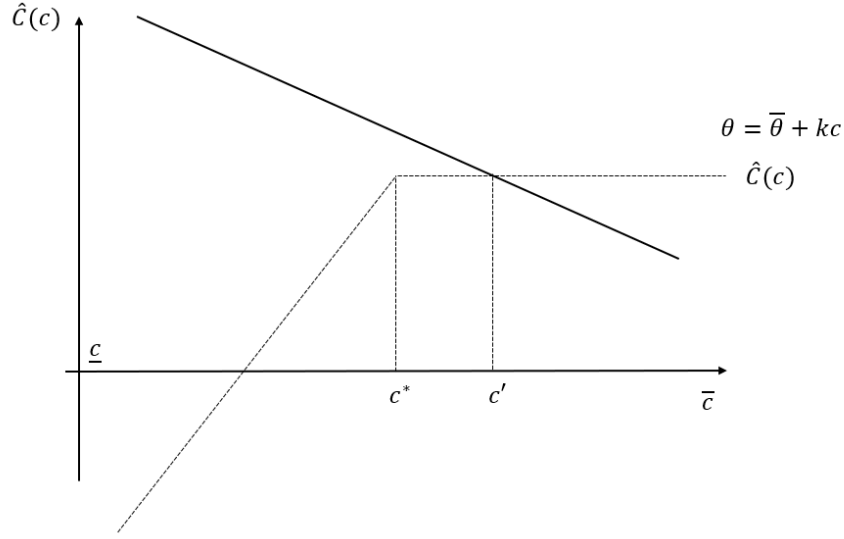


Figure 1: when $k < 0$, the threshold is distorted downwards to retain some of the high types as well.

Proof. Given the social welfare function in (28) the first-order condition for the abatement threshold, ie. carbon price is:

$$\begin{aligned} \frac{\partial W}{\partial c^*} = & (\gamma + D - c^* - \lambda T^*) f(c^*) + \frac{1}{k}(\gamma + \lambda(c^* - T^*))f(c') \\ & - (\gamma + \lambda(c^* - T^*))f(c^*) + \int_{c^*}^{c'} \lambda f(c) dc = 0 \end{aligned}$$

and the first-order condition for the base transfer T^* :

$$\frac{\partial W}{\partial T^*} = \int_{\underline{c}}^{c^*} (-\lambda)f(c)dc - \frac{1}{k}(\gamma + \lambda(c^* - T^*))f(c') + \int_{c^*}^{c'} (-\lambda)f(c)dc = 0$$

From which we get the optimal $\{c^*, T^*\}$:

$$c^* = \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{F(c^*)}{f(c^*)} < \frac{D}{1+\lambda} \quad (29)$$

and

$$\begin{aligned}
T^* &= \frac{\gamma}{\lambda} + c^* - \frac{kF(c')}{f(c')} \\
&= \frac{\gamma}{\lambda} + \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{F(c^*)}{f(c^*)} - \frac{kF(c')}{f(c')}
\end{aligned}$$

Since we have found an optimal solution that is a two-part tariff, we can conclude by Lemmas (1) & (2) that the resulting mechanism is the optimal IC mechanism. ■

3.2 $k > 1$: Strong positive association, where high-cost firms are the most immobile

When $k > 1$, then the situation is reversed from before. The high-cost firms now face the highest relocation costs, or are the most immobile. We assume that $c'' > \underline{c}$, so that we have an interior solution for the optimization, and conjecture that the partition of types in this case is the following: types $c < c''$ relocate, types $c'' \leq c \leq c^*$ stay and abate, and types $c > c^*$ stay without abating. Since with our conjecture, the only relocation constraint that is relevant is (24), that pins down the marginal relocating type below the abatement threshold c^* . The social welfare function to be optimized can be written as:

$$\max_{c^*, T^*} W = \int_{\frac{\bar{\theta} + T^*}{1-k}}^{c^*} (\gamma + (D - c) - \lambda T^*) f(c) dc + \int_{c^*}^{\bar{c}} (\gamma + \lambda(c^* - T^*)) f(c) dc \quad (30)$$

We summarize the optimal mechanism in this case in the following proposition.

Proposition 2. *Whenever $k > 1$, the high-cost firms are the most immobile, i.e. face the highest relocation costs, and the abatement threshold in the mechanism is distorted higher than first-best, to*

$$c^* = \frac{D}{1+\lambda} + \frac{\lambda}{1+\lambda} \frac{(1 - F(c^*))}{f(c^*)}$$

In this case we find that the threshold is now, surprisingly *distorted above* the first-best level, meaning that the country actually implements stricter regulation than would

be socially optimal. I illustrate this below in Figure 2.

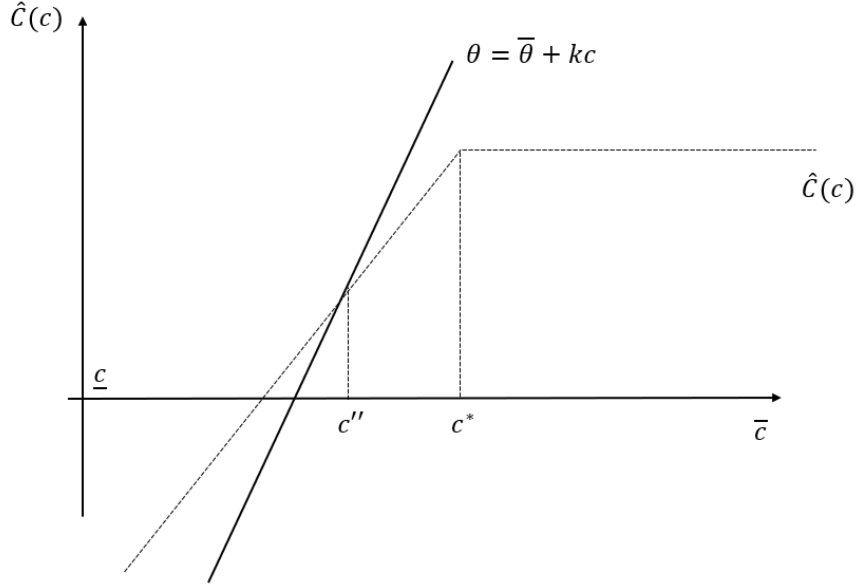


Figure 2: when $k > 1$, the threshold is distorted upwards to extract higher rents from the non-abating types above the threshold.

This distortion arises due to the fact that with this association regime, the high-cost firms face the highest relocation costs, and now the regulator has more leeway in designing the optimal c^* . Since the secondary role of c^* is a carbon tax for non-abaters, then increasing the threshold above the social optimum allows extracting more rents from the most immobile firms at the high end of the cost spectrum, while sacrificing relatively little abatement¹². This is similar to the upward distortion in Ahlvik and Liski (2021), which to my knowledge is a relatively non-standard result in the incentive regulation literature (although arising here perhaps in a slightly mechanistic way).

Proof. The first-order condition for c^* is:

$$\frac{\partial W}{\partial c^*} = (\gamma + D - c^* - \lambda T^*) f(c^*) - (\gamma + \lambda(c^* - T^*)) f(c^*) + \int_{c^*}^{\bar{c}} \lambda f(c) dc = 0$$

and

¹²Note that there exists also a corner solution to the optimization problem, where only non-abaters are retained.

$$c^* = \frac{D}{1+\lambda} + \frac{\lambda}{1+\lambda} \frac{(1-F(c^*))}{f(c^*)} > \frac{D}{1+\lambda} \quad (31)$$

the first-order condition for the base transfer is

$$\frac{\partial W}{\partial T^*} = -\frac{1}{1-k} \left(\gamma + D - \left(\frac{\bar{\theta} + T^*}{1-k} \right) - \lambda T^* \right) f(c'') + \int_{c''}^{c^*} (-\lambda) f(c) dc + \int_{c^*}^{\bar{c}} (-\lambda) f(c) dc = 0$$

from which

$$T^* = \frac{1-k}{\lambda(1-k)+1} (\gamma + D) + \frac{\lambda(1-k)^2(1-F(c''))}{(\lambda(1-k)+1)f(c'')} - \frac{\bar{\theta}}{\lambda(1-k)+1} \quad (32)$$

Since the solution is a two-part tariff, based on the previous lemmata we conclude that we have found the optimal mechanism for $k > 1$. ■

3.3 $0 < k < 1$: Mild positive association, both the highest-cost and lowest-cost firms are immobile

In the intermediate case **when** $0 < k < 1$ we have two relocation constraints that affect the partition of types, so both constraints (25) and (22) are relevant, pinning down both the lower and higher marginal relocating type. We assume that both $c' < \bar{c}$ and $c'' > \underline{c}$ so that we have an interior solution to the optimization problem. We conjecture that the partition in this case is the following: types $c \leq c'' = \frac{\bar{\theta} + T^*}{1-k}$ stay and abate, types in the middle, i.e. $c'' < c < c' := -\left(\frac{\bar{\theta} + T^* - c^*}{k}\right)$ leave, and types $c > c'$ stay without abating. Therefore the social welfare function is:

$$\max_{c^*, T^*} W = \int_{\underline{c}}^{\frac{\bar{\theta} + T^*}{1-k}} (\gamma + (D - c) - \lambda T^*) f(c) dc + \int_{-\left(\frac{\bar{\theta} + T^* - c^*}{k}\right)}^{\bar{c}} (\gamma + \lambda(c^* - T^*)) f(c) dc \quad (33)$$

Proposition 3. *Whenever $0 < k < 1$, both the lowest-cost and highest-cost firms are the most immobile, and the abatement threshold in the mechanism under global externalities lies between the thresholds of the previous mechanisms:*

$$c^* \in \left(\frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{F(c^*)}{f(c^*)}, \frac{D}{1+\lambda} + \frac{\lambda}{1+\lambda} \frac{(1-F(c^*))}{f(c^*)} \right)$$

Proof. Relegated to the Appendix for readability. ■

When the association between the abatement and relocation costs is mildly positive, we have a situation in which there are two marginal relocating types. One below the abatement threshold (c''), and one above (c'). In this case, surprisingly, the threshold is set above the marginal relocating type (c''), even though the types in (c'', c') all relocate. The reason this happens is that the threshold c^* is now used in its secondary role as a carbon price to tax the staying non-abaters optimally¹³. I note that interestingly, the first-best abatement threshold $c^{\text{FB}} = \frac{D}{1+\lambda}$ is contained in the interval, implying that in this association regime, the optimal mechanism even under asymmetric information may implement the first-best level of abatement.

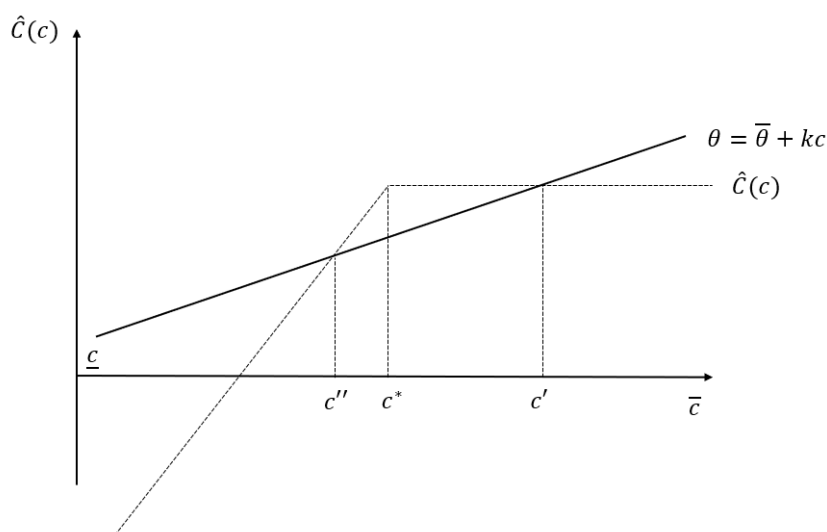


Figure 3: The optimal mechanism exploits the threshold c^* to tax the high types, when $0 < k < 1$.

¹³I focus only on the interior solution in this proposition, the corner solutions to this problem are relegated to the Appendix. One such corner solution is one that always obtains in this association regime and where the regulator retains all types. This is done by designing the compliance $\hat{C}(c)$ such that the kink lies on the linear outside option in the graph in Figure 3.

In each of the association regimes considered, the higher the shadow costs of public funds, the larger the distortions away from the first.best solution are. A somewhat less savory, but regardless immediate consequence of my chosen streamlined model is that whenever public funds have no shadow costs (in social welfare terms), the entire problem of asymmetric information disappears altogether.

4 Extensions to the leakage model

In this section I extend the previous leakage model by allowing the domestic regulator (country) to make *cross-border transfers*, i.e. to pay the relocating firms for their abatement abroad. When optimally designed, I show that this addition essentially does away with one major source of distortion in the previous leakage mechanisms, namely the fact that in some association regimes the regulator is losing firms which are efficient enough to yield positive social welfare *if* they could be incentivized to abate. Cross-border transfers now allow the regulator to create these incentives, and I show that in each association regime the regulator has a uniquely optimal cross-border transfer, which differs from the domestic carbon price, creating carbon price dispersion.

I extend the basic model considered previously by now allowing the regulator to pay the relocating firms a price $p_F \geq 0$ for their *voluntary* abatement abroad, or in other words we allow the regulator to make *cross-border transfers*. Therefore the regulatory mechanism, previously consisting only of an abatement threshold, or domestic carbon price c^* , and a base transfer or subsidy, T^* now becomes a triple, which we define as:

$$\mathcal{M} = \{c^*, p_F, T^*\}. \quad (34)$$

Where $p_F \geq 0$ is the price paid to a relocating firm in exchange for their voluntary abatement abroad, which I will call *foreign carbon price*. We assume that abatement by relocating firms is perfectly verifiable by the home country at zero cost to alleviate any monitoring concerns. The relocating firm will *voluntarily abate* abroad, iff $c \leq p_F$, since the firm can always reject the offer and simply relocate without accepting payment p_F . The relocation constraints for the firms are now different, since p_F enters them directly. To see how, let us express the relocation constraint of a firm, given an IC mechanism $\mathcal{M} = \{c^*, p_F, T^*\}$. As before, a type c -firm relocates if

$$\underbrace{\hat{C}(c)}_{\text{Domestic compliance}} > \bar{\theta} + kc + \underbrace{(c - p_f)}_{\text{Compensation for abatement abroad}} \quad (35)$$

and abates abroad when $p_f \geq c$. Furthermore, since relocating firms can choose to not accept the transfer and just relocate with no abatement, that implies that a moving firm of type c would reject any $p_f < c$, because that would entail a loss should they abate abroad. I derive here the uniquely optimal foreign carbon prices in the positive and negative association regimes, since they offer the cleanest version of the intuition behind these results¹⁴.

4.1 Strong positive association

In the association regimes we have considered, the strongest incentive for the regulator to set this price is when they lose the most efficient abaters. The association parameter that induces this scenario is strong positive association, i.e. when $k > 1$. Moreover, we note that securing abatement from abroad is only relevant for the regulator when the externality is global¹⁵. Therefore I begin with analyzing *strong positive association*, where $k > 1$.

In this regime, the low-cost firms are the most mobile, and the equilibrium mechanism in Proposition (2) and the associated Figure 2 illustrates the optimal domestic carbon price and the base transfer rate. Since the relocating firms are now the most efficient abaters, and since we are looking at interior solutions to the optimization problem, the regulator will always wish to offer $p_f > c$, which will secure a measure $F(p_f)$ of foreign abatement from the relocating firms. In this association regime, the following proposition characterizes the uniquely optimal foreign carbon price:

Proposition 4. *The uniquely optimal foreign carbon price is characterized by the minimum of the monopsony price and the marginal relocating type:*

¹⁴Full computations and results in the intermediate association regime are available by request.

¹⁵With local externalities, the relocating firms yield the maximum possible surplus obtainable from them (γ notwithstanding) at no additional cost to the regulator. Therefore purely local externalities will have the regulator either not offer this price, or alternatively, set $p_f = 0$.

$$p_F = \min \left(c'', \frac{D}{1+\lambda} - \frac{F(p_F)}{f(p_F)} \right), \quad \text{where } c'' := \frac{\bar{\theta} + T^*}{1-k}$$

Proof. When $k > 1$, the moving firms are ones with *low* abatement costs, so the regulator may wish to offer a positive price p_F to secure their abatement even in the alternative location. In this case, we conjecture that the partition is such that types $c \leq p_F$ move, *but abate abroad*, types $p_F < c \leq c''$ move and don't abate, types $c'' < c \leq c^*$ stay and abate, and types above c^* stay without abating, so the social welfare function is:

$$W = \int_{\underline{c}}^{p_F} (D - (1+\lambda)p_F) f(c) dc + \int_{\frac{\bar{\theta} + T^*}{1-k}}^{c^*} (\gamma + (D - c) - \lambda T^*) f(c) dc + \int_{c^*}^{\bar{c}} (\gamma + \lambda(c^* - T^*)) f(c) dc$$

where the first term is the extra surplus attained from abatement abroad. Taking the first-order conditions with respect to $\{T^*, c^*\}$ we find that they are the same as in case (ii) in Proposition 2, therefore the domestic mechanism is unchanged. The first-order condition with regards to the foreign price p_F is

$$\begin{aligned} \frac{\partial W}{\partial p_F} = 0 &\Leftrightarrow (D - (1+\lambda)p_F) f(p_F) - (1+\lambda)F(p_F) = 0 \Leftrightarrow \\ p_F &= \frac{D}{1+\lambda} - \frac{F(p_F)}{f(p_F)} \end{aligned}$$

Whenever $p_F < c''$, the regulator buys abatement from the moving firms as well. However, since the marginal relocating firm is of type $c'' = \frac{\bar{\theta} + T^*}{1-k}$ the regulator may be able to do even better than p_F . To see why, note that by simply setting $p_F = c''$, the regulator is buying the entire measure $[\underline{c}, c'']$ of reductions, and moreover every firm in this interval strictly prefers this scheme to one where they relocate without abatement. Therefore, the optimal global mechanism in this case consists of:

$$\left\{ c^*, T^*, \min \left(c'', \frac{D}{1+\lambda} - \frac{F(p_F)}{f(p_F)} \right) \right\} \quad (36)$$

where c^* and T^* are the same as in Proposition 2 (ii). Since the compensation for

abatement abroad is decreasing in the firm's type, and the marginal type abating abroad is pinned down by (36) the relocation constraints are unaffected in the domestic mechanism. ■

The addition of a foreign carbon price to the set of instruments available to the regulator essentially fixes the tradeoff facing the regulator in the simple leakage model; they are losing the most efficient firms that would yield the most social surplus when abating, which happens due to the strong positive association between abatement costs c and relocation costs θ . By adding in an instrument for the regulator that allows them to purchase the abatement from relocating firms, the regulator is able to reverse some of the carbon leakage, while still keeping the same measure of firms in-country. The base transfer rate is the same as in the simple leakage model (since the first-order condition is the same). This is due to the fact that the regulator has no need to adjust the base compensation or subsidies given to staying firms, since the relocation constraints are unaffected in this association regime. So instead, the regulator simply buys the abatement from the relocating firms at a price given above.

4.2 Negative association

Next let us consider the case of negative association, when $k < 0$. In this case, the situation is reversed from before, and the high-cost firms are the most mobile and relocate. Unlike in the previous association regime, the relocating firms now lie *above* the domestic abatement threshold, i.e. do not abate if they stay. Therefore, for the regulator to offer a positive foreign carbon price p_F , the relocating firms must not be too inefficient (in a sense I will make precise shortly), so that there exists a positive measure of firms that would yield a surplus if they abate abroad and are compensated for it by the linear scheme in (35).

Now, since the welfare gain available from any firm relocating and abating abroad for a price p_F is:

$$D - (1 + \lambda)p_F$$

It follows that the maximal price that the regulator is willing to pay is $p_F^{\text{MAX}} = \frac{D}{1+\lambda}$, which coincides with the first-best abatement threshold c^{FB} . We can then note that

whenever $c' = -\left(\frac{\bar{\theta} + T^* - c^*}{k}\right) > \frac{D}{1+\lambda}$, i.e. the marginal moving type lies above the first-best abatement threshold, the regulator is unwilling to offer the relocating firms a price they would require to abate abroad, because such a price would entail a strict welfare loss for the regulator, and therefore we assume

Assumption 2. *The marginal relocating type lies below the first-best abatement threshold, i.e.*

$$c' < c^{\text{FB}} = \frac{D}{1+\lambda}, \text{ where } c' := -\left(\frac{\bar{\theta} + T^* - c^*}{k}\right)$$

This assumption guarantees that there exists a positive measure of relocating high-cost firms that could still be incentivized to abate abroad at such a cost as to yield strictly positive social welfare for the domestic regulator¹⁶. Under this assumption, the optimal foreign carbon price, characterized below is strictly higher than the domestic carbon price c^* .

Proposition 5. *When the low-cost firms are the most immobile, and high-cost firms relocate, i.e. when $k < 0$ and moreover, whenever assumption (2) holds, the equilibrium mechanism is:*

$$\begin{aligned} c^* &= \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{F(c^*)}{f(c^*)} \\ p_F &= \frac{D}{1+\lambda} - \frac{1}{f(p_F)} [F(p_F) - F(c^\Delta)] \\ T^* &= \frac{\gamma}{\lambda} + c^* + \frac{1+\lambda}{\lambda} \left[\frac{F(c^\Delta)}{f(c^\Delta)} - \frac{F(p_F)}{f(p_F)} \right] + \frac{(1+k)F(c^\Delta)}{f(c^\Delta)} \end{aligned}$$

with the new induced marginal relocating type

$$c^\Delta = \frac{1}{k+1} (p_F + c^* - T^* - \bar{\theta}). \quad (37)$$

Comparing the equilibrium mechanism here to the one outlined in Proposition (1), we find that the domestic carbon price is the same than in the fully local mechanism

¹⁶If this assumption does not hold, then you may think that the regulator either offers a foreign carbon price of $p_F = 0$, or does not offer this price. An offer of $p_F = 0$ incentivizes at best a firm of type \underline{c} to abate, recalling that $\underline{c} \geq 0$.

derived previously but now the base transfer rate is different. This distortion in the transfer occurs since now the regulator can afford to pay less to firms it retains, since there exists a positive measure of relocating firms it can buy valuable abatement from, without having to worry about subsidizing their outside options.

Proof. First, we must establish the new marginal type, since the optimal choice of p_F will affect the relocation constraints of the staying non-abating firms. In particular, the old marginal type c' that is indifferent between staying and relocating, i.e.

$$c^* - T^* = kc' + \bar{\theta}$$

now has a strict preference to relocate, when offered $p_F > 0$ since

$$c^* - T^* > kc' + \bar{\theta} + \underbrace{(c' - p_F)}_{<0}$$

And therefore the price p_F induces the marginal relocating type to shift to the left, to type c^Δ , defined by

$$c^* - T^* = kc^\Delta + \bar{\theta} + (c^\Delta - p_F) \Leftrightarrow c^\Delta = \frac{1}{k+1} (p_F + c^* - T^* - \bar{\theta}) \quad (38)$$

Now we can again form the social welfare function in a piecewise manner, with the following partition: $c \leq c^*$ stay and abate, $c^* < c < c^\Delta$ stay and don't abate, and types $c \in [c^\Delta, p_F]$ relocate and abate. Therefore the SWF is

$$W := \int_{\underline{c}}^{c^*} (\gamma + D - c - \lambda T^*) f(c) dc + \int_{c^*}^{c^\Delta} (\gamma + \lambda(c^* - T^*)) + \int_{c^\Delta}^{p_F} (D - (1 + \lambda)\hat{c}) f(c) dc \quad (39)$$

The first-order conditions are:

$$\begin{aligned} \frac{\partial W}{\partial c^*} = & (\gamma + D - c^* - \lambda T^*) f(c^*) + \frac{1}{k+1} (\gamma + \lambda(c^* - T^*)) f(c^\Delta) \\ & - (\gamma + \lambda(c^* - T^*)) f(c^*) + \lambda \int_{c^*}^{c^\Delta} f(c) dc \\ & - \frac{1}{k+1} (D - (1 + \lambda)p_F) f(c^\Delta) = 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial W}{\partial p_F} = & \frac{1}{k+1} (\gamma + \lambda(c^* - T^*)) f(c^\Delta) + (D - (1 + \lambda)p_F) f(p_F) \\ & - \frac{1}{k+1} (D - (1 + \lambda)p_F) f(p_F) - (1 + \lambda) \int_{c^\Delta}^{p_F} f(c) dc = 0 \end{aligned}$$

and

$$\begin{aligned} \frac{\partial W}{\partial T^*} = & -\lambda \int_{c^*}^{c^\Delta} f(c) dc - \frac{1}{k+1} (\gamma + \lambda(c^* - T^*)) f(c^\Delta) - \lambda \int_{c^*}^{c^\Delta} f(c) dc \\ & + \frac{1}{k+1} (D - (1 + \lambda)p_F) f(c^\Delta) = 0 \end{aligned}$$

Rewriting the condition for T^* , we get

$$\frac{1}{k+1} (\gamma + \lambda(c^* - T^*)) f(c^\Delta) - \frac{1}{k+1} (D - (1 + \lambda)p_F) f(c^\Delta) = -\lambda F(c^\Delta)$$

inserting this into the condition for p_F we have

$$\begin{aligned} -\lambda F(c^\Delta) + (D - (1 + \lambda)p_F) f(p_F) - (1 + \lambda) [F(p_F) - F(c^\Delta)] = 0 \\ p_F = \frac{D}{1 + \lambda} - \frac{1}{f(p_F)} [F(p_F) - F(c^\Delta)] \end{aligned}$$

And since $f(p_F) > 0$ and $F(p_F) - F(c^\Delta) > 0$, we have that $p_F < p_F^{\text{MAX}} = \frac{D}{1 + \lambda}$. Using the same substitution in the condition for c^* , we get

$$\frac{\partial W}{\partial c^*} = (D - (1 + \lambda)c^*) f(c^*) - \lambda F(c^\Delta) + \lambda \int_{c^*}^{c^\Delta} f(c) dc = 0 \Leftrightarrow$$

$$c^* = \frac{D}{1 + \lambda} - \frac{\lambda}{1 + \lambda} \frac{F(c^*)}{f(c^*)}$$

which is the same threshold than in the previous domestic mechanism without the foreign price. The base transfer, however takes some more algebra to tease out, so start with rewriting the condition for T^* :

$$-\frac{1}{k+1} (\gamma + \lambda(c^* - T^*)) + \frac{1}{k+1} (D - (1 + \lambda)p_F) = \frac{\lambda F(c^\Delta)}{f(c^\Delta)}$$

inserting p_F and simplifying yields

$$T^* = \frac{\gamma}{\lambda} + c^* + \frac{1 + \lambda}{\lambda} \left[\frac{F(c^\Delta)}{f(c^\Delta)} - \frac{F(p_F)}{f(p_F)} \right] + \frac{(1 + k)F(c^\Delta)}{f(c^\Delta)}$$

■

5 Conclusions and discussion

I confirm, in a streamlined and stylized model of optimal environmental regulation under informational asymmetry and firm relocation risk the main insights of Ahlvik and Liski (2021), i.e. that carbon leakage, or firm relocation does not indicate a failed regulatory policy. Indeed, as highlighted above, carbon leakage in the form of firm relocation is, in fact, an equilibrium outcome in the *optimal second-best mechanism*. My model posits firm relocation risk as a type-dependent outside option, affiliated with the abatement costs of the firm in either a positive or negative way. This affiliation plays a key role in optimal policy design, essentially setting up multiple different affiliation regimes where the regulator is facing carbon leakage as a result of their regulatory policy, either by inefficient firms (negative association), or by efficient firms (strong positive association). The optimal regulatory policy is distorted either below or above first-best, where the upwards distortion is novel.

In a wider context, the results of this model imply that optimal regulatory policies under relocation risk have additional distortions beyond the simple ones caused by incentive compatibility, with both the size and sign of those distortions depending on the affiliation between the firm's cost type and its outside option. Therefore, my model formalizes one argument for regulating different industries separately, as one could easily think that different industries may well have different affiliations and hence face differing relocation, or carbon leakage risk. Industry-specific policies arise also in Hoel (1996) and in Martin, Muûls, De Preux, and Wagner (2014), where industry-specific leakage risk is one of the motivating factors behind the model.

I extend the model by relaxing the assumption that the regulator cannot make cross-border transfers to the relocating firms. I show that (conditional) cross-border transfers rectify a major drawback of the simple leakage mechanisms discussed previously: the fact that the regulator is losing socially valuable firms and therefore also abatement. With cross-border transfers, the regulator is able to buy the otherwise leaked abatement directly from the relocating firms themselves, essentially outsourcing not only the firms, but their abatement as well. Moreover, I show that in any association regime, the foreign carbon price is unique and does not coincide with the domestic carbon price, being either lower or higher, depending on the affiliation regime¹⁷.

¹⁷A similar result obtains when considering exogenous regulatory policies set by the "pollution haven" country; the other country's policy works to the domestic regulator's advantage, since any policy having

6 APPENDIX

6.1 Mild positive association (proof of Proposition 3)

Proof. the first-order condition for the carbon price is

$$\frac{\partial W}{\partial c^*} = -\frac{1}{k} (\gamma + \lambda(c^* - T^*)) f(c') + \int_{c'}^{\bar{c}} \lambda f(c) dc = 0$$

and the first-order condition for the base transfer is

$$\begin{aligned} \frac{\partial W}{\partial T^*} &= \frac{1}{1-k} \left(\gamma + D - \left(\frac{\bar{\theta} + T^*}{1-k} \right) - \lambda T^* \right) f(c'') + \int_{\underline{c}}^{c''} (-\lambda) f(c) dc \\ &\quad + \frac{1}{k} (\gamma + \lambda(c^* - T^*)) f(c') + \int_{c'}^{\bar{c}} (-\lambda) f(c) dc = 0 \end{aligned}$$

which yield, after some manipulations

$$c^* = T^* - \frac{\gamma}{\lambda} + \frac{k(1 - F(c'))}{f(c')} \quad (40)$$

and

$$T^* = \frac{1-k}{1+\lambda(1-k)} (\gamma + D) - \frac{\lambda(1-k)^2 F(c'')}{(1+\lambda(1-k))f(c'')} - \frac{\bar{\theta}}{1+\lambda(1-k)} \quad (41)$$

From the first-order condition for c^* :

$$\begin{aligned} \frac{\partial W}{\partial c^*} &= -\frac{1}{k} \underbrace{(\gamma + \lambda(c^* - T^*))}_{:= -\Delta(c')} f(c') + \int_{c'}^{\bar{c}} \lambda f(c) dc = 0 \\ \frac{\partial W}{\partial c^*} = 0 &\Leftrightarrow \Delta(c') f(c') = k(1 - F(c')) \end{aligned}$$

And the foc for T^* :

some bite in the other country for relocating firms serves to decrease the outside option of firms, hence allowing the domestic regulator to cut back on its own compensation paid to the firms it does retain.

$$\begin{aligned}
\frac{\partial W}{\partial T^*} &= \frac{1}{1-k} \underbrace{(\gamma + D - c'' - \lambda T^*)}_{:= -\Delta(c'')} f(c'') + \int_{\underline{c}}^{c''} (-\lambda) f(c) dc \\
&\quad + \underbrace{\frac{1}{k} (\gamma + \lambda(c^* - T^*)) f(c') + \int_{c'}^{\bar{c}} (-\lambda) f(c) dc}_{=\partial W/\partial c^*} = 0 \\
\frac{\partial W}{\partial T^*} = 0 &\Leftrightarrow \Delta(c'') f(c'') = -(1-k)\lambda F(c'')
\end{aligned}$$

Then, noting that $\Delta(c'') = \Delta(c') - (D - c'' + \lambda c^*)$ we can write

$$\begin{aligned}
\Delta(c'') f(c'') = -(1-k)\lambda F(c'') &\Leftrightarrow \Delta(c'') = -\frac{(1-k)\lambda F(c'')}{f(c'')} \\
\Delta(c') - (D - c'' + \lambda c^*) &= -\frac{(1-k)\lambda F(c'')}{f(c'')} \\
\frac{k\lambda(1-F(c'))}{f(c')} + D - c'' - \lambda c^* &= \frac{(1-k)\lambda F(c'')}{f(c'')} \\
D - c'' - \lambda c^* &= \frac{(1-k)\lambda F(c'')}{f(c'')} - \frac{k\lambda(1-F(c'))}{f(c')}
\end{aligned}$$

And, since by assumption $c'' = \left(\frac{\bar{\theta} + T^*}{1-k}\right) < c^*$, it follows that

$$D - (1-\lambda)c^* < \frac{(1-k)\lambda F(c'')}{f(c'')} - \frac{k\lambda(1-F(c'))}{f(c')}$$

And taking a look at the RHS, we have

$$\frac{\lambda F(c'')}{f(c'')} - k\lambda \left(\frac{F(c')}{f(c')} - \frac{F(c'')}{f(c'')} + \frac{1}{f(c')} \right)$$

Where we know that $\frac{F(c')}{f(c')} - \frac{F(c'')}{f(c'')} \geq 0$, by regularity of $F(\cdot)$ in c , and that $\frac{1}{f(c')} > 0$. Therefore the bracketed term is strictly positive and hence the second term strictly negative. Regularity of $F(\cdot)$ again implies that $\frac{\lambda F(c'')}{f(c'')} \leq \frac{\lambda F(c^*)}{f(c^*)}$, and so we have that

$$D - (1-\lambda)c^* < \frac{\lambda F(c^*)}{f(c^*)} \Leftrightarrow c^* > \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{F(c^*)}{f(c^*)}$$

meaning that the threshold in this intermediate case is strictly above the threshold in case (i). We have again arrived at a two-part tariff as our optimal solution, so by our previous lemmata we conclude this to be the optimal IC mechanism in this case. ■

6.2 Corner solutions in the basic leakage model

If we have a corner solution in any of the cases, i.e. if $c' \geq \bar{c}$ when $k < 0$, $c'' \leq \underline{c}$ when $k > 1$, or $c' \geq \bar{c}$ and $c'' \leq \underline{c}$ when $0 < k < 1$, the problem reduces to designing the optimal mechanism while no subset of types relocates. The objective function in every case and hence first-order condition for the threshold c^* is therefore

$$W = \int_{\underline{c}}^{c^*} (\gamma + (D - c) - \lambda T^*) f(c) dc + \int_{c^*}^{\bar{c}} (\gamma + \lambda(c^* - T^*)) f(c) dc$$

$$\frac{\partial W}{\partial c^*} = (\gamma + D - c^* - \lambda T^*) f(c^*) - (\gamma + \lambda(c^* - T^*)) f(c^*) + (1 - F(c^*)) \lambda = 0$$

which yields $c^* = \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{(1-F(c^*))}{f(c^*)}$. The first-order conditions for the transfer T^* coincide as well in each case, and therefore we have the same optimal mechanism on offer in any regime of k .

Moreover, whenever there is weak positive association, i.e. $0 < k < 1$, one corner solution then is for the regulator to exploit the optimality of the two-part tariff characterization of all IC mechanisms by designing the mechanism such that $\theta \geq \hat{C}(c)$ for all c . This is done by requiring that the only indifferent marginal type is type c^* , therefore pinning down the transfer T^* , since:

$$\begin{aligned} c^* - T^* &= \theta \\ &= \bar{\theta} + kc^* \\ T^*(c^*) &= (1 - k)c^* - \bar{\theta} \end{aligned}$$

with this in hand, we can now express the SWF and optimize with regard to c^* alone, leading to:

$$\begin{aligned}
W &= \int_{\underline{c}}^{c^*} (\gamma + (D - c) - \lambda T^*) f(c) dc + \int_{c^*}^{\bar{c}} (\gamma + \lambda(c^* - T^*)) f(c) dc \\
&= \int_{\underline{c}}^{c^*} (\gamma + (D - c) - \lambda [(1 - k)c^* - \bar{\theta}]) f(c) dc + \int_{c^*}^{\bar{c}} (\gamma + \lambda c^* - \lambda [(1 - k)c^* - \bar{\theta}]) f(c) dc
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial W}{\partial c^*} &= (\gamma + D - c^* - \lambda(1 - k)c^* + \lambda\bar{\theta}) - \lambda(1 - k) \int_{\underline{c}}^{c^*} f(c^*) dc \\
&\quad - (\gamma + \lambda c^* - \lambda(1 - k)c^* + \lambda\bar{\theta}) f(c^*) + \lambda k \int_{c^*}^{\bar{c}} f(c) dc = 0
\end{aligned}$$

from which we get

$$c^* = \frac{D}{1 + \lambda} - \frac{\lambda}{1 + \lambda} \frac{1}{f(c^*)} [k - F(c^*)]$$

and the transfer is then computed as previously. As before with the actual leakage model, we immediately get a result that the threshold c^* may lie below or above the first-best level of $\frac{D}{1 + \lambda}$. Interestingly, if it happens that

$$k = F(c^*)$$

this mechanism actually implements the first-best level of abatement.

Optimal mechanisms for $k = 0$, $k = 1$:

If we have no association between the abatement and relocation costs, i.e. $k = 0$, then the relocation cost θ is constant at $\bar{\theta}$ for all firm types. Moreover, we restrict ourselves to relocation costs $\bar{\theta} \geq 0$, so that our model does not become non-sensical. The optimization problem now reduces to a standard mechanism design problem with a fixed outside option, since type-dependence is now ruled out. In this case, Lemmas 1 and 2 still imply the uniquely optimal IC mechanism to be a two-part tariff. Now, however,

firms relocate when

$$\hat{C}(c) \geq \bar{\theta}$$

Proposition 6. *When $k = 0$, the optimal mechanism is:*

$$c^* = \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{F(c^*)}{f(c^*)}$$

$$T^* = c^* - \bar{\theta}$$

Proof. It is immediate that the regulator can retain the highest measure of firms (i.e. receive the highest welfare possible) by reducing the non-abating firms to their outside option payoff. This is now possible, since the outside option is a constant. This then implies that we can solve for the base transfer using the compliance cost of a non-abating firm:

$$c^* - T^* = \bar{\theta} \Leftrightarrow T^* = c^* - \bar{\theta}$$

This scheme retains all the types, and therefore the social welfare function is:

$$W(c^*) = \int_{\underline{c}}^{c^*} (\gamma + (D - c) - \lambda(c^* - \bar{\theta})) f(c) dc + \int_{c^*}^{\bar{c}} (\gamma + \lambda\bar{\theta}) f(c) dc$$

The first-order condition for the threshold is:

$$\frac{\partial W}{\partial c^*} = (\gamma + D - c^* - \lambda(c^* - \bar{\theta})) f(c^*) - \lambda F(c^*) - (\gamma + \lambda\bar{\theta}) f(c^*) = 0$$

from which $c^* = \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda} \frac{F(c^*)}{f(c^*)}$. Since the mechanism we have derived is a two-part tariff, by our previous lemmata we conclude it to be uniquely optimal in this case. ■

When $k = 1$, then the relocation cost is an affine function of the abatement cost of the firm, i.e. $\theta = c + \bar{\theta}$. Now, unlike in the previous case, here the regulator will optimally use a mechanism that reduces *the abating types* to their outside option payoffs, while leaving information rents to the non-abaters. Due to our assumption of linear abatement costs, where the cost of abating a single unit of pollution for a firm is just c , we see immediately that the relocation cost θ now has the same slope as the abatement cost for any firm type. This now means that the regulator can essentially use a mechanism that lines these two functions up (by clever design of the base transfer, T^*), and therefore guarantee that every abater will be reduced to their outside option payoff.

for this to be the case, we must then have that the compliance for the abating firms leaves them at their outside option payoff, i.e.

$$c - T^* = c + \bar{\theta} \quad (42)$$

from which we get that $T^* = -\bar{\theta}$. At an interior solution, the objective function becomes (after inserting T^* from before):

$$W = \int_{\underline{c}}^{c^*} (\gamma + (D - c) + \lambda \bar{\theta}) f(c) dc + \int_{c^*}^{\bar{c}} (\gamma + \lambda(c^* - \bar{\theta})) f(c) dc$$

The first-order condition for the threshold gives us the same upwards-distorted threshold as in the $k > 1$ association regime, i.e. we have that:

$$c^* = \frac{D}{1 + \lambda} + \frac{\lambda}{1 + \lambda} \frac{(1 - F(c^*))}{f(c^*)}$$

However, we have one other candidate for an equilibrium mechanism in this case. As previously mentioned, when $k = 1$ the regulator reduces the abating types to their outside option payoffs. This also means that the regulator has at their disposal an IC mechanism, where *all the types abate*. To see how this is constructed, note that if the regulator sets a threshold of $\hat{c} = \bar{c}$, then every firm will face a compliance cost as in (42), since all of them are abating if they stay. As in the previous section, this induced compliance also pins down the base transfer, such that $T^* = -\bar{\theta}$. Incentive compatibility is trivially satisfied, since no deviation by any firm type to any report will change neither

their allocation nor their transfer. Therefore we have a second candidate mechanism in this case, one consisting of

$$\hat{\mathcal{M}} = \{\bar{c}, -\bar{\theta}\} \quad (43)$$

The expected welfare from this mechanism is

$$\begin{aligned} \hat{W} &= \int_{\underline{c}}^{\bar{c}} (\gamma + D - c + \lambda\bar{\theta}) f(c) dc \\ &= D + \gamma + \lambda\bar{\theta} - \underbrace{\int_{\underline{c}}^{\bar{c}} cf(c) dc}_{=E[c]} \end{aligned}$$

While the first mechanism yields an expected welfare of:

$$\begin{aligned} W &= \int_{\underline{c}}^{\bar{c}} (\gamma + \lambda\bar{\theta}) f(c) dc + \int_{\underline{c}}^{c^*} (D - c) f(c) dc + \int_{c^*}^{\bar{c}} \lambda c^* f(c) dc \\ &= \gamma + \lambda\bar{\theta} + \int_{\underline{c}}^{c^*} (D - c) f(c) dc + \int_{c^*}^{\bar{c}} \lambda c^* f(c) dc \end{aligned}$$

computing the difference in welfare between the two, we have

$$\begin{aligned} \hat{W} - W &= D - \left(\int_{\underline{c}}^{c^*} (D - c) f(c) dc + \int_{c^*}^{\bar{c}} \lambda c^* f(c) dc \right) - \int_{\underline{c}}^{\bar{c}} cf(c) dc \\ &= D - \int_{\underline{c}}^{c^*} (D - 2c) f(c) dc - \int_{c^*}^{\bar{c}} (\lambda c^* - c) f(c) dc \end{aligned}$$

and from this expression, it is possible to define a range for the parameters (D, λ) such that the integrands are negative a.e., and hence the difference $\hat{W} - W > 0$ almost everywhere. In this parameter range then, the *optimal* mechanism under $k = 0$ is therefore the one specified in (43). Outside this range, the optimal mechanism is

$$\mathcal{M} = \left\{ \frac{D}{1+\lambda} + \frac{\lambda}{1+\lambda} \frac{(1-F(c^*))}{f(c^*)}, -\bar{\theta} \right\} \quad (44)$$

Bibliography

- AGUIRRE, I., AND A. BEITIA (2017): "Modelling countervailing incentives in adverse selection models: A synthesis," *Economic Modelling*, 62, 82–89.
- AHLVIK, L., AND M. LISKI (2017): "Carbon leakage: a mechanism design approach," *Working paper*.
- (2021): "Global Externalities, Local Policies and Firm Selection," *The Journal of the European Economic Association* (forthcoming).
- AHLVIK, L., M. LISKI, AND R. MARTIN (2017): "The association between the vulnerability score and compliance costs," *Mimeo*.
- ARMSTRONG, M., AND J.-C. ROCHET (1999): "Multi-dimensional screening: A user's guide," *European Economic Review*, 43(4-6), 959–979.
- BALIGA, S., AND E. MASKIN (2003): "Mechanism design for the environment," in *Handbook of environmental economics*, vol. 1, pp. 305–324. Elsevier.
- BARON, D. P. (1985): "Regulation of prices and pollution under incomplete information," *Journal of Public Economics*, 28(2), 211–231.
- DASGUPTA, P., P. HAMMOND, AND E. MASKIN (1980): "On imperfect information and optimal pollution control," *The Review of Economic Studies*, 47(5), 857–860.
- DECHEZLEPRÊTRE, A., AND M. SATO (2017): "The Impacts of Environmental Regulations on Competitiveness," *Review of Environmental Economics and Policy*, 11(2), 183–206.
- EDERINGTON, J., A. LEVINSON, AND J. MINIER (2005): "Footloose and pollution-free," *Review of Economics and Statistics*, 87(1), 92–99.
- ESKELAND, G. S., AND A. E. HARRISON (2003): "Moving to greener pastures? Multinationals and the pollution haven hypothesis," *Journal of development economics*, 70(1), 1–23.
- FOWLIE, M., AND M. REGUANT (2018): "Challenges in the measurement of leakage risk," in *AEA Papers and Proceedings*, vol. 108, pp. 124–29.
- HOEL, M. (1996): "Should a carbon tax be differentiated across sectors?," *Journal of public economics*, 59(1), 17–32.

- KIM, J.-C., AND K.-B. CHANG (1993): "An optimal tax/subsidy for output and pollution control under asymmetric information in oligopoly markets," *Journal of Regulatory Economics*, 5(2), 183–197.
- LAFFONT, J.-J., AND J. TIROLE (1993): *A theory of incentives in procurement and regulation*. MIT Press.
- (1996): "Pollution permits and compliance strategies," *Journal of Public Economics*, 62(1), 85–125.
- LEVINSON, A., AND M. S. TAYLOR (2008): "Unmasking the pollution haven effect," *International economic review*, 49(1), 223–254.
- LEWIS, T. R., AND D. E. SAPPINGTON (1989): "Countervailing incentives in agency problems," *Journal of economic theory*, 49(2), 294–313.
- (1995): "Using markets to allocate pollution permits and other scarce resource rights under limited information," *Journal of Public Economics*, 57(3), 431–455.
- MARTIN, R., M. MUÛLS, L. B. DE PREUX, AND U. WAGNER (2014): "Industry compensation under relocation risk: A firm-level analysis of the EU emissions trading scheme," *The American Economic Review*, 104(8), 2482–2508.
- MISCH, F., AND P. WINGENDER (2021): "Revisiting Carbon Leakage," *IMF Working Paper 21/07*.
- MONTERO, J.-P. (2008): "A simple auction mechanism for the optimal allocation of the commons," *The American Economic Review*, 98(1), 496–518.
- NAEGELE, H., AND A. ZAKLAN (2019): "Does the EU ETS cause carbon leakage in European manufacturing?," *Journal of Environmental Economics and Management*, 93, 125–147.
- ROCHET, J.-C., AND L. A. STOLE (2002): "Nonlinear pricing with random participation," *The Review of Economic Studies*, 69(1), 277–311.
- VENMANS, F., J. ELLIS, AND D. NACHTIGALL (2020): "Carbon pricing and competitiveness: are they at odds?," *Climate Policy*, 20(9), 1070–1091.
- VERDE, S. F. (2020): "The impact of the EU emissions trading system on competitiveness and carbon leakage: the econometric evidence," *Journal of economic surveys*, 34(2), 320–343.

VISLIE, J. (2000): "Environmental regulation under asymmetric information with type-dependent outside option," *Memorandum No. 2000, 18, University of Oslo*.

YU, B., Q. ZHAO, AND Y.-M. WEI (2021): "Review of carbon leakage under regionally differentiated climate policies," *Science of The Total Environment*, 782, 146765.

Bank of Finland Research Discussion Papers 2023

ISSN 1456-6184, online

- 1/2023 Gonalo Faria – Fabio Verona
Forecast combination in the frequency domain
ISBN 978-952-323-430-7, online
- 2/2023 Saara Tuuli
Who funds zombie firms: banks or non-banks?
ISBN 978-952-323-431-4, online
- 3/2023 Olli-Matti Laine – Matias Pihlajamaa
Asymmetric effects of conventional and unconventional monetary policy when rates are low
ISBN 978-952-323-432-1, online
- 4/2023 Mingze Gao – Iftekhar Hasan – Buhui Qiu – Eliza Wu
Lone (Loan) Wolf Pack Risk
ISBN 978-952-323-433-8, online
- 5/2023 Mathias Drehmann – Mikael Juselius – Anton Korinek
Long-term debt propagation and real reversals
ISBN 978-952-323-434-5, online
- 6/2023 Simon Kwan – Mauricio Ulate – Ville Voutilainen
The transmission of negative nominal interest rates in Finland
ISBN 978-952-323-435-2, online
- 7/2023 Iftekhar Hasan – Haekwon Lee – Buhui Qiu – Anthony Saunders
Climate-related disclosure commitment of the lenders, credit rationing, and borrower environmental performance
ISBN 978-952-323-436-9, online
- 8/2023 Sanggyup Choi – Jaehun Jeong – Dohyeon Park – Donghoon Yoo
News or animal spirits? Consumer confidence and economic activity: Redux
ISBN 978-952-323-437-6, online
- 9/2023 Francesco D'Acunto – Daniel Hoang – Maritta Paloviita – Michael Weber
Cognitive Constraints and Economic Incentives
ISBN 978-952-323-438-3, online
- 10/2023 Szilard Benk – Max Gillman
Identifying money and inflation expectation shocks on real oil prices
ISBN 978-952-323-442-0, online
- 11/2023 Zuzana Fungáčov – Eeva Kerola – Olli-Matti Laine
Monetary policy transmission below zero
ISBN 978-952-323-443-7, online
- 12/2023 Alexander Duck – Fabio Verona
Monetary policy rules: model uncertainty meets design limits
ISBN 978-952-323-445-1, online
- 13/2023 Olli-Matti Laine – Jaakko Nelimarkka
Assessing targeted longer-term refinancing operations: identification through search intensity
ISBN 978-952-323-446-8, online
- 14/2023 Bonnie Buchanan – Hanna Silvola – Emilia Vahamaa
Sustainability and private investors
ISBN 978-952-323-448-2, online
- 15/2023 Tero Hokkanen
Optimal carbon leakage
ISBN 978-952-323-449-9, online